

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2002, 11:00:11 ; search time 29.95 Seconds
(without alignments)
467.289 Million cell updates/sec

Title: US-09-863-063-2

perfect score: 126
Sequence: 1 AQSVPPEDIQTKPDKIVFN.....EWFGQDDGMVRKKNLPIEYNP 126

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 11107396 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_032802:*

```

1: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb/AA2001.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	15	11.9	16	18	AAW31324		Dictyocaulus vivip
2	9	7.1	10	18	AAW31325		Dictyocaulus vivip
3	8	6.3	8	18	AAW31323		Dictyocaulus vivip
4	8	6.3	20	18	AAW31322		Dictyocaulus vivip
5	8	6.3	454	14	AAR33365		Sequence of purine
6	7	5.6	73	22	AAO01320		Human polypeptide
7	7	5.6	142	22	AM80296		Human protein SEQ
8	7	5.6	151	22	AO012550		Human polypeptide
9	7	5.6	154	22	AAG74054		Human colon cancer
10	7	5.6	155	21	AAG26775		Zea mays protein F
11	7	5.6	160	22	AM80302		Human protein SEQ

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	15	11.9	16	18	AAW31324		Dictyocaulus vivip
2	9	7.1	10	18	AAW31325		Dictyocaulus vivip
3	8	6.3	8	18	AAW31323		Dictyocaulus vivip
4	8	6.3	20	18	AAW31322		Dictyocaulus vivip
5	8	6.3	454	14	AAR33365		Sequence of purine
6	7	5.6	73	22	AAO01320		Human polypeptide
7	7	5.6	142	22	AM80296		Human protein SEQ
8	7	5.6	151	22	AO012550		Human polypeptide
9	7	5.6	154	22	AAG74054		Human colon cancer
10	7	5.6	155	21	AAG26775		Zea mays protein F
11	7	5.6	160	22	AM80302		Human protein SEQ

1	AAW31324	standard; peptide; 16 AA.
2	AB239402	Peptide #6119 enco
3	ABB24184	Peptide #6119 enco
4	AAM60082	Novel lung cancer
5	AAM72691	Novel human respir
6	AAM19685	Arabidopsis thalia
7	AAM32922	Arabidopsis thalia
8	AAU18940	Human brain expres
9	AAU17952	Human bone marrow
10	AAG08061	Human papillomavir
11	AAG61624	Human immune/haema
12	AAG89881	Human protein SEQ
13	AAY79881	Novel human diagno
14	AAR30176	Novel immunogen
15	AAY25777	Staphylococcus aur
16	AAY3632	Staphylococcus aur
17	AAY68778	Staphylococcus aur
18	AAY74054	Staphylococcus aur
19	AAM88344	Staphylococcus aur
20	AAM56623	Staphylococcus aur
21	AAM78344	Staphylococcus aur
22	AAM79328	Staphylococcus aur
23	AAM79344	Staphylococcus aur
24	AAB33632	Staphylococcus aur
25	AAB33632	Staphylococcus aur
26	AAB33632	Staphylococcus aur
27	AAB33632	Staphylococcus aur
28	AAB33632	Staphylococcus aur
29	AAB33632	Staphylococcus aur
30	AAB33632	Staphylococcus aur
31	AAB33632	Staphylococcus aur
32	AAB33632	Staphylococcus aur
33	AAB33632	Staphylococcus aur
34	AAB33632	Staphylococcus aur
35	AAB33632	Staphylococcus aur
36	AAB33632	Staphylococcus aur
37	AAB33632	Staphylococcus aur
38	AAB33632	Staphylococcus aur
39	AAB33632	Staphylococcus aur
40	AAB33632	Staphylococcus aur
41	AAB33632	Staphylococcus aur
42	AAB33632	Staphylococcus aur
43	AAB33632	Staphylococcus aur
44	AAB33632	Staphylococcus aur
45	AAB33632	Staphylococcus aur

RESULT 1
AAW31324 standard; peptide; 16 AA.
ID AAW31324 (first entry)
XX DE Dictyocaulus viviparus DV18 peptide fragment 4.
XX KW Immunogenic protein; DV18; lungworm; vaccine; immunity;
KW dictyocauliasis; cattle; immunoassay.
XX OS Dictyocaulus viviparus.
XX PN EP785253-Al.
XX PD 23-JUL-1997.
XX PF 27-DEC-1996; 96EP-0120947.
XX PR 19-JAN-1996; 96DE-4001754.
(FARH) HOECHST AG.
PI Hofmann J, Schmid K;
XX DR WPI; 1997-365928/34.
XX PT Immunogenic Dictyocaulus viviparus lungworm protein - for use in vaccines and immunoassays
XX PS Example 10; Page 6; 17pp; German.

CC This sequence represents a peptide fragment of the immunogenic protein DV18 isolated from adult lungworms. The N-terminal of this amino acid can be joined to the amino acids Asp, Asn, Ser, Gly or Arg. DV18 can be used in vaccines for immunising cattle against dictyocauliasis and in an ELISA immunoassay for determining DV18-specific antibodies in the blood of cattle.

XX Sequence 16 AA:

Query Match 11.9%; Score 15; DB 18; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 FRREWFQDGGMVRRK 118
| | | | | | | | | | | |
Db 2 frrewfqdggmvrrk 16

RESULT 2
ID AAW31325 standard; peptide: 10 AA.
XX AAW31325;

AC

XX DT 13-MAR-1998 (first entry)

DE Dictyocaulus viviparus DV18 peptide fragment 5.

XX Immunogenic protein; DV18; lungworm; vaccine; immunity;

KW dictyocauliasis; cattle; immunoassay.

XX DE Dictyocaulus viviparus DV18 peptide fragment 5.

XX Immunogenic protein; DV18; lungworm; vaccine; immunity;

KW dictyocauliasis; cattle; immunoassay.

XX DE Dictyocaulus viviparus DV18 peptide fragment 5.

XX Immunogenic protein; DV18; lungworm; vaccine; immunity;

KW dictyocauliasis; cattle; immunoassay.

XX DE Dictyocaulus viviparus DV18 peptide fragment 5.

XX Immunogenic protein; DV18; lungworm; vaccine; immunity;

KW dictyocauliasis; cattle; immunoassay.

XX DE Dictyocaulus viviparus DV18 peptide fragment 5.

XX Immunogenic protein; DV18; lungworm; vaccine; immunity;

KW dictyocauliasis; cattle; immunoassay.

XX DE Dictyocaulus viviparus DV18 peptide fragment 5.

XX Immunogenic protein; DV18; lungworm; vaccine; immunity;

KW dictyocauliasis; cattle; immunoassay.

XX DE Dictyocaulus viviparus DV18 peptide fragment 5.

XX Immunogenic protein; DV18; lungworm; vaccine; immunity;

KW dictyocauliasis; cattle; immunoassay.

XX DE Dictyocaulus viviparus DV18 peptide fragment 5.

XX Immunogenic protein; DV18; lungworm; vaccine; immunity;

KW dictyocauliasis; cattle; immunoassay.

XX DE Dictyocaulus viviparus DV18 peptide fragment 5.

XX Immunogenic protein; DV18; lungworm; vaccine; immunity;

KW dictyocauliasis; cattle; immunoassay.

XX DE Dictyocaulus viviparus DV18 peptide fragment 5.

XX Immunogenic protein; DV18; lungworm; vaccine; immunity;

KW dictyocauliasis; cattle; immunoassay.

XX DE Dictyocaulus viviparus DV18 peptide fragment 5.

XX Immunogenic protein; DV18; lungworm; vaccine; immunity;

KW dictyocauliasis; cattle; immunoassay.

XX DE Dictyocaulus viviparus DV18 peptide fragment 5.

XX Immunogenic protein; DV18; lungworm; vaccine; immunity;

KW dictyocauliasis; cattle; immunoassay.

XX DE Dictyocaulus viviparus DV18 peptide fragment 5.

XX Immunogenic protein; DV18; lungworm; vaccine; immunity;

KW dictyocauliasis; cattle; immunoassay.

XX DE Dictyocaulus viviparus DV18 peptide fragment 5.

XX Immunogenic protein; DV18; lungworm; vaccine; immunity;

KW dictyocauliasis; cattle; immunoassay.

XX DE Dictyocaulus viviparus DV18 peptide fragment 5.

XX Immunogenic protein; DV18; lungworm; vaccine; immunity;

CC AAW31323 standard; peptide: 8 AA.
ID AAW31323
XX
AC AAW31323;
XX DT 13-MAR-1998 (first entry)

XX DE Dictyocaulus viviparus DV18 peptide fragment 3.

XX KW Immunogenic protein; DV18; lungworm; vaccine; immunity;

XX KW dictyocauliasis; cattle; immunoassay.

XX OS Dictyocaulus viviparus.

XX PN EP785253-A1.

XX PD 23-JUL-1997.

XX PR 19-JAN-1996;

XX PA (FARH) HOECHST AG.

XX PI Hofmann J, Schmid K;

XX DR WPI; 1997-365928/34.

XX PT WPI; 1997-365928/34.

XX PT Vaccines and immunoassays

XX XX Example 10; Page 6; 17pp; German.

XX PS EP785253-A1.

XX DR 23-JUL-1997.

XX PR 27-DEC-1996;

XX PA (FARH) HOECHST AG.

XX PI Hofmann J, Schmid K;

XX DR WPI; 1997-365928/34.

XX PT WPI; 1997-365928/34.

XX PT Vaccines and immunoassays

XX XX Example 10; Page 6; 17pp; German.

XX PS (FARH) HOECHST AG.

XX PI Hofmann J, Schmid K;

XX DR WPI; 1997-365928/34.

XX PR 19-JAN-1996;

XX PA 96DE-4001754.

XX XX Sequence 8 AA;

XX PA (FARH) HOECHST AG.

XX PI Hofmann J, Schmid K;

XX DR WPI; 1997-365928/34.

XX PR 27-DEC-1996;

XX PA (FARH) HOECHST AG.

XX PI Hofmann J, Schmid K;

XX DR WPI; 1997-365928/34.

XX PT WPI; 1997-365928/34.

XX PT Vaccines and immunoassays

XX XX Example 10; Page 6; 17pp; German.

XX PS (FARH) HOECHST AG.

XX PI Hofmann J, Schmid K;

XX DR WPI; 1997-365928/34.

XX PR 19-JAN-1996;

XX PA 96DE-4001754.

XX XX Sequence 8 AA;

XX PA (FARH) HOECHST AG.

XX PI Hofmann J, Schmid K;

XX DR WPI; 1997-365928/34.

XX PR 27-DEC-1996;

XX PA (FARH) HOECHST AG.

XX PI Hofmann J, Schmid K;

XX DR WPI; 1997-365928/34.

XX PT WPI; 1997-365928/34.

XX PT Vaccines and immunoassays

XX XX Example 10; Page 6; 17pp; German.

XX PS (FARH) HOECHST AG.

XX PI Hofmann J, Schmid K;

XX DR WPI; 1997-365928/34.

XX PR 19-JAN-1996;

XX PA 96DE-4001754.

XX XX Sequence 8 AA;

XX PA (FARH) HOECHST AG.

XX PI Hofmann J, Schmid K;

XX DR WPI; 1997-365928/34.

XX PR 19-JAN-1996;

XX PA 96DE-4001754.

XX XX Sequence 8 AA;

XX PA (FARH) HOECHST AG.

XX PI Hofmann J, Schmid K;

XX DR WPI; 1997-365928/34.

XX PR 19-JAN-1996;

XX PA 96DE-4001754.

XX XX Sequence 8 AA;

XX PA (FARH) HOECHST AG.

XX PI Hofmann J, Schmid K;

XX DR WPI; 1997-365928/34.

XX PR 19-JAN-1996;

XX PA 96DE-4001754.

XX XX Sequence 8 AA;

XX PA (FARH) HOECHST AG.

XX PI Hofmann J, Schmid K;

XX DR WPI; 1997-365928/34.

XX PR 19-JAN-1996;

XX PA 96DE-4001754.

XX XX Sequence 8 AA;

XX PA (FARH) HOECHST AG.

XX PI Hofmann J, Schmid K;

XX DR WPI; 1997-365928/34.

XX PR 19-JAN-1996;

XX PA 96DE-4001754.

XX XX Sequence 8 AA;

XX PA (FARH) HOECHST AG.

XX PI Hofmann J, Schmid K;

XX DR WPI; 1997-365928/34.

XX PR 19-JAN-1996;

XX PA 96DE-4001754.

XX XX Sequence 8 AA;

XX PA (FARH) HOECHST AG.

XX PI Hofmann J, Schmid K;

XX DR WPI; 1997-365928/34.

XX PR 19-JAN-1996;

XX PA 96DE-4001754.

XX XX Sequence 8 AA;

XX PA (FARH) HOECHST AG.

PD 23-JUL-1997.
 XX 96EP-0120947.
 PF 27-DEC-1996; 96EP-0120947.
 XX 19-JAN-1996; 96DE-4 001754.
 PR (FARH) HOECHST AG.
 XX
 PA Hofmann J, Schmid K;
 XX DR; 1997-365928/34.
 XX PT Immunogenic Dictyocaulus viviparus lungworm protein - for use in vaccines and immunoassays
 XX PS Example 10; Page 6; 17pp; German.
 XX This sequence represents a peptide fragment of the immunogenic protein DV18 isolated from adult lungworms. The N-terminal of this amino acid can be joined to the amino acids Asp, Asn, Ser, Gly or Arg. DV18 can be used in vaccines for immunising cattle against dictyocauliasis and in an ELISA immunoassay for determining DV18-specific antibodies in the blood of cattle.
 XX SQ Sequence 20 AA;
 XX Query Match 6.3%; Score 8; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 60 GVLDPKEA 67
 Db 9 gvldpkea 16

RESULT 5
 AAR333365 standard; Protein; 454 AA.
 ID AAR333365 standard; Protein; 454 AA.
 XX AC AAR333365;
 XX DT 30-JUN-1993 (first entry)
 XX DE Sequence of purine-rich repeat (GA repeat) binding protein (GABP) subunit alpha.
 XX KW GA binding protein; cis-regulatory element;
 XX KW VP16 mediated induction.
 XX OS Mus musculus.
 XX PN WO9304166-A.
 XX PD 04-MAR-1993.
 XX PF 17-AUG-1992; 92WO-US06748.
 XX PR 16-AUG-1991; 91US-0746032.
 XX PA (CARN-) CARNEGIE INST WASHINGTON.
 XX PI Lamarco KL, Mc Knight SL, Thompson CC;
 XX DR WPI; 1993-093998/11.
 XX DR N-PSDB; AAQ37480.
 XX PT DNA encoding GA binding protein sub-unit - allows investigation of sub-unit sequence motif functions, for control of rapid cell division e.g. in cancer
 XX PT Disclosure; Fig 2A; 68pp; English.
 XX PS

CC A cis-regulatory element required for virion associated protein VP16 mediated induction of herpes simplex virus 1 (HSV1) immediate early (IE) genes consists of three imperfect repeats of the purine-rich hexanucleotide 5'-CGGAAR-3'. A protein complex capable of avid interaction with the purine-rich repeats (GA repeats) has been identified in soluble preparations of rat liver nucleic. This GA binding protein (GABP) consists of two separable subunits. Applicants have isolated cDNA clones encoding both subunits of GABP and have revealed that one (GABP alpha) is related to the Ets transforming protein, while the other (GABP beta) contains a series of 33-amino acid repeats related in sequence to a variety of proteins.
 XX SQ Sequence 454 AA;

Query Match 6.3%; Score 8; DB 14; Length 454;
 Best Local Similarity 100.0%; Pred. No. 4.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 31 IKVINSSA 38
 Db 284 ikvinssa 291

RESULT 6
 AAO01320
 ID AAO01320 standard; Protein; 73 AA.
 XX AC AAO01320;
 XX DT 06-NOV-2001 (first entry)
 XX DE Human polypeptide SEQ ID NO 15212.
 XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematoopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX OS Homo sapiens.
 XX PN WO200164835-A2.
 XX PD 07-SEP-2001.
 XX PF 26-FEB-2001; 2001WO-US04927.
 XX PR 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Drmanac RT;
 XX DR WPI; 2001-514838/56.
 DR N-PSDB; AA181251.
 XX PT Isolated nucleic acids and polypeptides, useful for preventing PT diagnosing and treating e.g. leukaemia, inflammation and immune PT disorders -
 XX Claim 20; SEQ ID NO 15212; 1399pp + Sequence Listing; English.
 XX
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and the encoded Proteins (AA000010-AA013910) that exhibit activity elating to CC cytokine, cell proliferation or cell differentiation or which may induce CC production of other cytokines in other cell populations. The CC polynucleotides and polypeptides are useful in gene therapy, vaccines or CC peptide therapy. The polypeptides have various cytokine-like activities, CC e.g. stem cell growth factor activity, haematoopoiesis regulating CC activity, tissue growth factor activity, immunomodulatory activity and CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic format directly from WIPO CC at ftp://wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 73 AA;

Query Match 5.6%; Score 7; DB 22; Length 73;
 Best Local Similarity 100.0%; Pred. No. 9.4%; Mismatches 0; Gaps 0;
 Matches 7; Conservative 0; Indels 0; Gaps 0;

Qy 97 PDGAARQ 103
 Db 24 pdgaarq 30

RESULT 7
 AAM80296 standard; Protein: 142 AA.
 XX ID AAM80296 standard; Protein: 151 AA.
 AC AAO12550;
 XX DT 06-NOV-2001 (first entry)
 DE Human protein SEQ ID NO 3948.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX OS Homo sapiens.
 XX PN WO200157190-A2.
 XX PR 09-AUG-2001.
 XX PF 05-FEB-2001; 2001WO-US04098.
 XX PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Weijhrman T, Goodrich R;
 XX DR WPI; 2001-476283/51.
 DR N-PSDB; AAK53429.

PS Claim 20; Page 6210-6211; 6221PP; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to CC cytokine, cell proliferation or cell differentiation or which may induce CC production of other cytokines in other cell populations. The CC polynucleotides and polypeptides are useful in gene therapy, vaccines or CC peptide therapy. The polypeptides have various cytokine-like activities, CC e.g. stem cell growth factor activity, haematoopoiesis regulating CC activity, tissue growth factor activity, immunomodulatory activity and CC activin/inhibin activity and may be useful in the diagnosis and/or CC treatment of cancer, leukaemia, nervous system disorders, arthritis and CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic format directly from WIPO CC at ftp://wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 142 AA;

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 CC (AAM80020) are omitted as the relevant pages from the sequence listing CC were missing at the time of publication.
 XX

Query Match 5.6%; Score 7; DB 22; Length 142;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 AVLLAVS 73
 Db 2 avllavs 8

RESULT 8
 AAO12550
 ID AAO12550 standard; Protein: 151 AA.
 XX AC AAO12550;
 XX DT 06-NOV-2001 (first entry)
 DE Human polypeptide SEQ ID NO 26442.
 XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX OS Homo sapiens.
 XX PN WO200164835-A2.
 XX PR 07-SEP-2001.
 XX PF 26-FEB-2001; 2001WO-US04927.
 XX PR 28-FEB-2000; 2000US-0515126.
 PR 18-MAY-2000; 2000US-0577409.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Drmanac RT;
 XX DR WPI; 2001-514838/56.
 DR N-PSDB; AAI92481.

XX PT Isolated nucleic acids and polypeptides, useful for preventing CC diagnosing and treating e.g. leukaemia, inflammation and immune CC disorders -
 XX PS Claim 20; SEQ ID NO 26442; 1399pp + Sequence Listing; English.
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and CC the encoded proteins (AA00010-AA013910) that exhibit activity relating to CC cytokine, cell proliferation or cell differentiation or which may induce CC production of other cytokines in other cell populations. The CC polynucleotides and polypeptides are useful in gene therapy, vaccines or CC peptide therapy. The polypeptides have various cytokine-like activities, CC e.g. stem cell growth factor activity, haematoopoiesis regulating CC activity, tissue growth factor activity, immunomodulatory activity and CC activin/inhibin activity and may be useful in the diagnosis and/or CC treatment of cancer, leukaemia, nervous system disorders, arthritis and CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic format directly from WIPO CC at ftp://wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 151 AA;

PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.	Query	97	Match	5.68;	Score 7;	DB 21;	Length 155;
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.	Best	99US-0161406.	Local	100.0%;	Pred.	No. 19;	
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.	Similarity	99US-0161359.			Mismatches	0;	
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.	Matches	7;	Conservative	0;			
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.	Qy	97	DGAAARQ	103			
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.	Db	10					
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.		11	pdgaarq	18			
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.	RESULT	11					
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.							
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.							
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.							
PR	23-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.							
PR	23-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.							
PR	21-JUN-1999;	99US-0139817.	PR	07-SEP-1999;	99US-0152363.							
PR	22-JUN-1999;	99US-0139899.	PR	10-SEP-1999;	99US-0151303.							
PR	23-JUN-1999;	99US-0140353.	PR	13-SEP-1999;	99US-0153758.							
PR	01-JUL-1999;	99US-0140354.	PR	15-SEP-1999;	99US-0154018.							
PR	24-JUN-1999;	99US-0140695.	PR	16-SEP-1999;	99US-0154039.							
PR	28-JUN-1999;	99US-0140823.	PR	20-SEP-1999;	99US-0154779.							
PR	29-JUN-1999;	99US-0140991.	PR	22-SEP-1999;	99US-0155139.							
PR	30-JUN-1999;	99US-0141287.	PR	23-SEP-1999;	99US-01553758.							
PR	01-JUL-1999;	99US-0141842.	PR	24-SEP-1999;	99US-0155659.							
PR	01-JUL-1999;	99US-0142154.	PR	28-SEP-1999;	99US-0156458.							
PR	02-JUL-1999;	99US-0142055.	PR	29-SEP-1999;	99US-0156596.							
PR	06-JUL-1999;	99US-0142390.	PR	04-OCT-1999;	99US-0157117.							
PR	08-JUL-1999;	99US-0142803.	PR	05-OCT-1999;	99US-0157753.							
PR	09-JUL-1999;	99US-0142920.	PR	06-OCT-1999;	99US-0157865.							
PR	12-JUL-1999;	99US-0142977.	PR	07-OCT-1999;	99US-0158029.							
PR	13-JUL-1999;	99US-0143542.	PR	08-OCT-1999;	99US-0158232.							
PR	14-JUL-1999;	99US-0143624.	PR	12-OCT-1999;	99US-0158369.							
PR	15-JUL-1999;	99US-0144005.	PR	13-OCT-1999;	99US-0159293.							
PR	16-JUL-1999;	99US-0144085.	PR	13-OCT-1999;	99US-0159294.							
PR	16-JUL-1999;	99US-0144086.	PR	13-OCT-1999;	99US-0159295.							
PR	19-JUL-1999;	99US-0144325.	PR	14-OCT-1999;	99US-0159329.							
PR	19-JUL-1999;	99US-0144331.	PR	14-OCT-1999;	99US-0159330.							
PR	19-JUL-1999;	99US-0144332.	PR	14-OCT-1999;	99US-0159331.							
PR	19-JUL-1999;	99US-0144333.	PR	14-OCT-1999;	99US-0159637.							
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159638.							
PR	19-JUL-1999;	99US-0144335.	PR	18-OCT-1999;	99US-0159884.							
PR	20-JUL-1999;	99US-0144352.	PR	21-OCT-1999;	99US-0160814.							
PR	20-JUL-1999;	99US-0144632.	PR	21-OCT-1999;	99US-0160815.							
PR	20-JUL-1999;	99US-0144884.	PR	21-OCT-1999;	99US-0160767.							
PR	21-JUL-1999;	99US-0144814.	PR	21-OCT-1999;	99US-0160770.							
PR	22-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160814.							
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.							
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160768.							
PR	22-JUL-1999;	99US-0145087.	PR	22-OCT-1999;	99US-0160981.							
PR	23-JUL-1999;	99US-0145089.	PR	22-OCT-1999;	99US-0160989.							
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.							
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.							
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.							
PR	28-JUL-1999;	99US-0145224.	PR	26-OCT-1999;	99US-0161359.							
PR	02-AUG-1999;	99US-0145276.	PR	26-OCT-1999;	99US-0161360.							
PR	02-AUG-1999;	99US-0146388.	PR	26-OCT-1999;	99US-0161361.							
PR	04-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.							
PR	04-AUG-1999;	99US-0147302.	PR	28-OCT-1999;	99US-0161922.							
PR	05-AUG-1999;	99US-0147192.	PR	29-OCT-1999;	99US-0162142.							
PR	05-AUG-1999;	99US-0147260.										
PR	06-AUG-1999;	99US-0147303.										
PR	06-AUG-1999;	99US-0147416.										
PR	09-AUG-1999;	99US-0147493.										
PR	09-AUG-1999;	99US-0147935.										
PR	10-AUG-1999;	99US-0148171.										
PR	11-AUG-1999;	99US-0148319.										
PR	12-AUG-1999;	99US-0148341.										
PR	13-AUG-1999;	99US-0148565.										

Query Match 5.68;
Best Local Similarity 100.0%;
Matches 7; Conservative 0;
Mismatches 0; Indels 0; Gaps 0;

Db 12 pdgaarq 18

AAM80302
ID AAM80302 standard; Protein; 160 AA.
XX
AC AAM80302;
XX DT 06-NOV-2001 (first entry)
XX DE Human protein SEQ ID NO 3960.
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX OS Homo sapiens.
XX PN WO200157190-A2.
XX PD 09-AUG-2001.
XX PF 05-FEB-2001; 2001WO-US04098.
XX PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX PA (HYSE-) HYSEQ INC.
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Weijhrman T, Goodrich R;
XX DR 2001-476283/51.
N-PSDB; AAK53435.
XX PS Claim 20; Page 477; 6221pp; English.
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX PS Claim 20; Page 477; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM8020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX SQ Sequence 160 AA;
SQ 5.6%; Score 7; DB 22; Length 160;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query Match 5.6%; Score 7; DB 22; Length 161;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 KRLGVDP 57
| | | | |
Db 120 krlgvdp 126
RESULT 12

RESULT 13
 AAG26274 ID AAG26274 standard; Protein; 175 AA.
 XX AC AAG26274;
 XX DT 17-OCT-2000 (first entry)
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 30670.
 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX OS Arabidopsis thaliana.
 XX PN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0301439.
 XX PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 24-JUN-1999; 99US-0140354.
 PR 28-JUN-1999; 99US-0140695.
 PR 29-JUN-1999; 99US-0140823.
 PR 30-JUN-1999; 99US-0140991.
 PR 01-JUL-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142300.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145192.
 PR 22-JUL-1999; 99US-0145145.
 PR 22-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 22-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

DT 18-OCT-2000 (first entry)
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 59019.
 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX OS Arabidopsis thaliana.
 XX PN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000;; 2000EP-0301439.
 XX PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 14-JUN-1999; 99US-0138847.
 PR 18-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.

Query Match 5.68; Score 7; DB 21; Length 175;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 AVILLAVS 73
 DB 54 avillavs 60

RESULT 14
 AAG46875 standard; Protein; 175 AA.
 XX AAG46875:
 AC XX

PR 18-JUN-1999; 99US-0139763. PR 27-AUG-1999; 99US-0151066.
 PR 21-JUN-1999; 99US-0139817. PR 27-AUG-1999; 99US-0151080.
 PR 22-JUN-1999; 99US-0139899. PR 30-AUG-1999; 99US-0151303.
 PR 23-JUN-1999; 99US-0140353. PR 31-AUG-1999; 99US-0151438.
 PR 23-JUN-1999; 99US-0140354. PR 01-SEP-1999; 99US-0151930.
 PR 24-JUN-1999; 99US-0140695. PR 07-SEP-1999; 99US-0152363.
 PR 28-JUN-1999; 99US-0140823. PR 10-SEP-1999; 99US-0153070.
 PR 29-JUN-1999; 99US-0140991. PR 13-SEP-1999; 99US-0153758.
 PR 30-JUN-1999; 99US-0141287. PR 15-SEP-1999; 99US-0154018.
 PR 01-JUL-1999; 99US-0141842. PR 16-SEP-1999; 99US-0154039.
 PR 09-JUL-1999; 99US-0142154. PR 20-SEP-1999; 99US-0154779.
 PR 02-JUL-1999; 99US-0142055. PR 22-SEP-1999; 99US-0155139.
 PR 06-JUL-1999; 99US-0142390. PR 23-SEP-1999; 99US-0155486.
 PR 08-JUL-1999; 99US-0142803. PR 24-SEP-1999; 99US-0155659.
 PR 09-JUL-1999; 99US-0142920. PR 28-SEP-1999; 99US-0156458.
 PR 12-JUL-1999; 99US-0142977. PR 29-SEP-1999; 99US-0156596.
 PR 13-JUL-1999; 99US-0143542. PR 04-OCT-1999; 99US-0157117.
 PR 14-JUL-1999; 99US-0143624. PR 23-OCT-1999; 99US-0157753.
 PR 15-JUL-1999; 99US-0144005. PR 05-OCT-1999; 99US-0157865.
 PR 16-JUL-1999; 99US-0144085. PR 06-OCT-1999; 99US-0158029.
 PR 16-JUL-1999; 99US-0144086. PR 07-OCT-1999; 99US-0158232.
 PR 19-JUL-1999; 99US-0144325. PR 08-OCT-1999; 99US-0158369.
 PR 19-JUL-1999; 99US-0144331. PR 09-OCT-1999; 99US-0159293.
 PR 19-JUL-1999; 99US-0144332. PR 13-OCT-1999; 99US-0159294.
 PR 19-JUL-1999; 99US-0144333. PR 13-OCT-1999; 99US-0159295.
 PR 19-JUL-1999; 99US-0144334. PR 14-OCT-1999; 99US-0159329.
 PR 19-JUL-1999; 99US-0144335. PR 14-OCT-1999; 99US-0159331.
 PR 20-JUL-1999; 99US-0144352. PR 14-OCT-1999; 99US-0159337.
 PR 20-JUL-1999; 99US-0144632. PR 14-OCT-1999; 99US-0159638.
 PR 20-JUL-1999; 99US-0144884. PR 18-OCT-1999; 99US-0159584.
 PR 21-JUL-1999; 99US-0144814. PR 21-OCT-1999; 99US-0160741.
 PR 21-JUL-1999; 99US-0145086. PR 21-OCT-1999; 99US-0160767.
 PR 21-JUL-1999; 99US-0145088. PR 21-OCT-1999; 99US-0160768.
 PR 22-JUL-1999; 99US-0145085. PR 21-OCT-1999; 99US-0160770.
 PR 22-JUL-1999; 99US-0145087. PR 21-OCT-1999; 99US-0160814.
 PR 22-JUL-1999; 99US-0145089. PR 21-OCT-1999; 99US-0161404.
 PR 22-JUL-1999; 99US-0145192. PR 21-OCT-1999; 99US-0161405.
 PR 23-JUL-1999; 99US-0145145. PR 25-OCT-1999; 99US-0161406.
 PR 23-JUL-1999; 99US-0145218. PR 25-OCT-1999; 99US-0161359.
 PR 23-JUL-1999; 99US-0145224. PR 26-OCT-1999; 99US-0161360.
 PR 26-JUL-1999; 99US-0145276. PR 26-OCT-1999; 99US-0161361.
 PR 27-JUL-1999; 99US-0145913. PR 28-OCT-1999; 99US-0161920.
 PR 27-JUL-1999; 99US-0145918. PR 28-OCT-1999; 99US-0161921.
 PR 27-JUL-1999; 99US-0145919. PR 28-OCT-1999; 99US-0161922.
 PR 28-JUL-1999; 99US-0145951. PR 28-OCT-1999; 99US-0161923.
 PR 02-AUG-1999; 99US-0146386. PR 29-OCT-1999; 99US-0162142.
 PR 02-AUG-1999; 99US-0146388. PR 29-OCT-1999; 99US-0162142.
 PR 02-AUG-1999; 99US-0146389. PR 29-OCT-1999; 99US-0162142.
 PR 03-AUG-1999; 99US-0147038. PR 29-OCT-1999; 99US-0162142.
 PR 04-AUG-1999; 99US-0147204. PR 29-OCT-1999; 99US-0162142.
 PR 04-AUG-1999; 99US-0147302. PR 29-OCT-1999; 99US-0162142.
 PR 05-AUG-1999; 99US-0147192. PR 29-OCT-1999; 99US-0162142.
 PR 06-AUG-1999; 99US-0147260. PR 29-OCT-1999; 99US-0162142.
 PR 06-AUG-1999; 99US-0147303. PR 29-OCT-1999; 99US-0162142.
 PR 06-AUG-1999; 99US-0147416. PR 29-OCT-1999; 99US-0162142.
 PR 09-AUG-1999; 99US-0147493. PR 29-OCT-1999; 99US-0162142.
 PR 09-AUG-1999; 99US-0147935. PR 29-OCT-1999; 99US-0162142.
 PR 10-AUG-1999; 99US-0148171. PR 29-OCT-1999; 99US-0162142.
 PR 11-AUG-1999; 99US-0148319. PR 29-OCT-1999; 99US-0162142.
 PR 12-AUG-1999; 99US-0148341. PR 29-OCT-1999; 99US-0162142.
 PR 13-AUG-1999; 99US-0148565. PR 29-OCT-1999; 99US-0162142.
 PR 13-AUG-1999; 99US-0148684. PR 29-OCT-1999; 99US-0162142.
 PR 16-AUG-1999; 99US-0149368. PR 29-OCT-1999; 99US-0162142.
 PR 17-AUG-1999; 99US-0149175. PR 29-OCT-1999; 99US-0162142.
 PR 18-AUG-1999; 99US-0149426. PR 29-OCT-1999; 99US-0162142.
 PR 20-AUG-1999; 99US-0149722. PR 29-OCT-1999; 99US-0162142.
 PR 20-AUG-1999; 99US-0149723. PR 29-OCT-1999; 99US-0162142.
 PR 20-AUG-1999; 99US-0149929. PR 29-OCT-1999; 99US-0162142.
 PR 23-AUG-1999; 99US-0149902. PR 29-OCT-1999; 99US-0162142.
 PR 23-AUG-1999; 99US-0149930. PR 29-OCT-1999; 99US-0162142.
 PR 25-AUG-1999; 99US-0150566. PR 29-OCT-1999; 99US-0162142.
 PR 26-AUG-1999; 99US-0150884. PR 29-OCT-1999; 99US-0162142.

Query Match 5.6%; Score 7; DB 21; Length 175;
 Best Local Similarity 100.0%; pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0;
 Indels 0; Gaps 0;

RESULT 15
 ID AAG89942 standard; Protein; 178 AA.
 AC AAG89942;
 Db 54 avllavs 60

QY 67 AVLLAVS 73
 DE C glutamicum protein fragment SEQ ID NO: 3696.
 XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis;
 XX

DT 26-SEP-2001 (first entry)
 XX
 DE
 AC
 XX

OS Corynebacterium glutamicum.
 XX EP1108790-A2.
 PN 20-JUN-2001.
 PD XX
 PF 18-DEC-2000; 2000EP-0127688.
 XX PR 16-DEC-1999; 99JP-0377484.
 PR 07-APR-2000; 2000JP-0159162.
 PR 03-AUG-2000; 2000JP-0280988.
 XX PA (KYOWA) KYOWA HAKKO KOGYO KK.
 XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX DR WPI; 2001-376931/40.
 DR N-PSDB; AAH65161.
 XX PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analysing
 PT expression profile or pattern of a gene and identifying homologous gene
 PT
 XX PS Claim 17; SEQ ID NO: 3696; 246pp + Sequence Listing; English.
 XX CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and
 CC analysing the expression profile or expression pattern of a gene derived
 CC from Coryneform bacterium, and identifying a homologue of a gene derived
 CC from coryneform bacterium. Coryneform bacteria are useful for producing
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described
 CC in the exemplification of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.
 XX SQ Sequence 178 AA;
 XX

Query Match 5.6%; Score 7; DB 22; Length 178;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 67 AVLLAVS 73
 Db 71 avllavs 77

Search completed: September 24, 2002, 11:02:54
 Job time: 163 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2002, 11:00:36 ; Search time 12.89 Seconds
(without alignments)
238.761 Million cell updates/sec

Title: US-09-863-063-2

perfect score: 126 Sequence: 1 AQSVPPGDIQTQPGTKIVFN.....EWFQGDGMVRKRNLPPIEYNP 126

Scoring table: OLIGO Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:/*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:/*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:/*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:/*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:/*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:/*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	18.3	127	4 US-08-786-455B-14	Sequence 14, Appl
2	10	7.9	14	4 US-08-786-455B-4	Sequence 4, Appl
3	9	7.1	9	4 US-08-786-455B-5	Sequence 5, Appl
4	8	6.3	20	4 US-08-786-455B-2	Sequence 2, Appl
5	7	5.6	1066	2 US-08-308-818-1	Sequence 1, Appl
6	7	5.6	1551	4 US-09-425-665-2	Sequence 2, Appl
7	7	5.6	1551	4 US-09-685-668-2	Sequence 2, Appl
8	6	4.8	7	4 US-08-786-455B-3	Sequence 3, Appl
9	6	4.8	34	3 US-08-658-136-34	Sequence 34, Appl
10	6	4.8	71	4 US-08-928-213B-23	Sequence 23, Appl
11	6	4.8	106	2 US-08-785-065-11	Sequence 11, Appl
12	6	4.8	128	3 US-08-906-769-143	Sequence 143, Appl
13	6	4.8	128	3 US-08-906-616-143	Sequence 143, Appl
14	6	4.8	128	4 US-08-639-075A-143	Sequence 143, Appl
15	6	4.8	128	4 US-09-012-431-143	Sequence 143, Appl
16	6	4.8	128	4 US-09-012-692-143	Sequence 143, Appl
17	6	4.8	128	4 US-08-906-613-143	Sequence 143, Appl
18	6	4.8	209	4 US-09-196-293-15	Sequence 15, Appl
19	6	4.8	210	1 US-08-158-353-3	Sequence 3, Appl
20	6	4.8	210	4 US-08-209-603E-15	Sequence 15, Appl
21	6	4.8	210	4 US-08-235-836C-30	Sequence 30, Appl
22	6	4.8	252	4 US-09-199-637A-176	Sequence 176, Appl
23	6	4.8	264	1 US-08-107-042-2	Sequence 2, Appl
24	6	4.8	273	1 US-08-215-928A-2	Sequence 2, Appl
25	6	4.8	273	1 US-08-446-920-2	Sequence 9, Appl
26	6	4.8	273	4 US-08-928-213B-9	Sequence 9, Appl
27	6	4.8	309	2 US-08-701-191A-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-08-786-455B-14
; Sequence 14, Application US/08786455B
; Patent No. 6193971

; GENERAL INFORMATION:
; APPLICANT: HOFMANN, Joachim
; APPLICANT: SCHMID, Karlheinrich
; TITLE OF INVENTION: DICTYOCARUS VIVIPARUS ANTIGEN FOR
; TITLE OF INVENTION: DIAGNOSING LUNGWORM INFESTATION AND FOR VACCINATION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/786,455B
; FILING DATE: 21-JAN-1997
; CLASSIFICATION: 536
; PRIORITY APPLICATION NUMBER: DE 196 01 754 .8
; FILING DATE: 19-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: GRANADOS, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 18748/327
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1.27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-786-455B-14

Query Match 18.3%; Score 23; DB 4; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.5e-16;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 QFRREWFGDGMVRRKNLPIEYN 125
 | | | | | | | | | | | | | | | | | |
 Db 103 QFRREWFGDGMVRRKNLPIEYN 125

RESULT 2
 US-08-786-455B-4
 ; Sequence 4, Application US/08786455B
 ; Patent No. 6193971
 ; GENERAL INFORMATION:
 ; APPLICANT: HOFMANN, Joachim
 ; SCHMID, Karlheinrich
 ; TITLE OF INVENTION: DICTYOCaulus viviparus antigen for
 ; diagnosing lungworm infestation and for vaccination
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/786,455B
 FILING DATE: 21-JAN-1997
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE 196 01 754 .8
 FILING DATE: 19-JAN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: GRANADOS, Patricia D.
 REGISTRATION NUMBER: 33,683
 REFERENCE/DOCKET NUMBER: 18748/327
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)672-5300
 TELEFAX: (202)672-5399
 TELEX: 904136

INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-786-455B-5

Query Match 7.1%; Score 9; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.7e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE 196 01 754 .8
 FILING DATE: 19-JAN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: GRANADOS, Patricia D.
 REGISTRATION NUMBER: 33,683
 REFERENCE/DOCKET NUMBER: 18748/327
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)672-5300
 TELEFAX: (202)672-5399
 TELEX: 904136

INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 14 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-786-455B-4

Query Match 7.9%; Score 10; DB 4; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.00063;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 FRREWFGDGC 113
 | | | | | | | | | | | | | | | | | |
 Db 1 FRREWFGDGC 10

RESULT 3
 US-08-786-455B-5
 ; Sequence 5, Application US/08786455B
 ; Patent No. 6193971
 ; GENERAL INFORMATION:
 ; APPLICANT: HOFMANN, Joachim
 ; SCHMID, Karlheinrich
 ; TITLE OF INVENTION: DICTYOCaulus viviparus antigen for
 ; diagnosing lungworm infestation and for vaccination
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500

CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20007-5109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/786,455B
 FILING DATE: 21-JAN-1997
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DE 196 01 754 .8
 FILING DATE: 19-JAN-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: GRANADOS, Patricia D.

```

REGISTRATION NUMBER: 33, 683
REFERENCE/DOCKET NUMBER: 18748/327
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: Linear

US-08-786-455B-2

Query Match 5.6%; Score 7; DB 2; Length 1066;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 PGTKIVF 19
Db 504 PGTKIVF 510

RESULT 6
US-09-425-665-2
Sequence 2, Application US/09425665
; Patent No. 6174705
; GENERAL INFORMATION:
; APPLICANT: Nicholas, Richard O.
; TITLE OF INVENTION: ARO1
; FILE REFERENCE: GM10245
; CURRENT APPLICATION NUMBER: US/09/425,665
; CURRENT FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1551
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-425-665-2

Query Match 5.6%; Score 7; DB 4; Length 1551;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 VLLAVSC 74
Db 1453 VLLAVSC 1459

RESULT 7
US-09-685-668-2
Sequence 2, Application US/09685668
; Patent No. 6346405
; GENERAL INFORMATION:
; APPLICANT: Green, Michael R
; TITLE OF INVENTION: A No. 5847077el Fungal Multisubunit Protein
; TITLE OF INVENTION: Complex Critical for Expression of Fungal Proteins
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: US
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DCS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/308,818
; FILING DATE: 19-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, S. Peter
; REGISTRATION NUMBER: 25,351
; REFERENCE/DOCKET NUMBER: 0342/0A404
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-52707700
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1066 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: S. cerevisiae
IMMEDIATE SOURCE:
CLONE: TAF-145
US-08-308-818-1

Query Match 5.6%; Score 7; DB 4; Length 1551;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 VLLAVSC 74
Db 1453 VLLAVSC 1459

RESULT 8
US-08-786-455B-3
Sequence 3, Application US/08786455B
; Patent No. 6193971

```

GENERAL INFORMATION:
 APPLICANT: HOFFMANN, Joachim
 APPLICANT: SCHMID, Karlheinrich
 TITLE OF INVENTION: DICTYOCaulus VIVIPARUS ANTIGEN FOR
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 3000 K Street, N.W., Suite 500
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20007-5109

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/658,136
 FILING DATE:

CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: LASSEN, ELIZABETH
 REGISTRATION NUMBER: 31,845
 REFERENCE/DOCKET NUMBER: GEN4-17.8

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 508-872-8400
 TELEFAX: 508-872-5415
 INFORMATION FOR SEQ ID NO: 34:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 34 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-658-136-34

Query Match 4.88; Score 6; DB 3; Length 34;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QSVPNG 7
 Db 18 QSVPNG 23

RESULT 10
 US-08-928-213B-23
 ; Sequence 23, Application US/08928213B
 ; Patent No. 6238905
 ; GENERAL INFORMATION:
 ; APPLICANT: McHenry, Charles S.
 ; Seville, Mark
 ; Cull, Millard G.
 ; TITLE OF INVENTION: NOVEL THERMOPHILIC POLYMERASE III
 ; NUMBER OF SEQUENCES: HOLOENZYME
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MEDLEN & CARROLL, LLP
 ; STREET: 220 Montgomery Street, Suite 2200
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94104

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,213B
 FILING DATE: 12-Sep-1997
 CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: MacKnight, Kamrin T.
 REGISTRATION NUMBER: 38,230
 REFERENCE/DOCKET NUMBER: ENZYCO-02550
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-705-8410
 TELEFAX: 415-397-8338
 INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 71 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: not relevant

Query Match 4.8%; Score 6; DB 4; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.7e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 LPTEYN 125
 Db 1 LPTEYN 6

RESULT 9
 US-08-658-136-34
 ; Sequence 34, Application US/08658136
 ; Patent No. 6071717
 ; GENERAL INFORMATION:
 ; APPLICANT: KLINGER, KATHERINE W
 ; APPLICANT: LANDES, GREGORY M
 ; APPLICANT: BURN, TIMOTHY C
 ; APPLICANT: CONNORS, TIMOTHY D
 ; APPLICANT: DACKOWSKI, WILLIAM
 ; APPLICANT: GERMINO, GREGORY
 ; APPLICANT: QIAN, FENG

TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
 NUMBER OF SEQUENCES: 58
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENZYME CORPORATION
 STREET: ONE MOUNTAIN ROAD
 CITY: FRAMINGHAM
 STATE: MASSACHUSETTS
 COUNTRY: USA
 ZIP: 01701

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,213B
 FILING DATE: 12-Sep-1997
 CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: MacKnight, Kamrin T.
 REGISTRATION NUMBER: 38,230
 REFERENCE/DOCKET NUMBER: ENZYCO-02550
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-705-8410
 TELEFAX: 415-397-8338
 INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 71 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: not relevant

;

;

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-08-928-213B-23

Query Match 4.8%; Score 6; DB 4; Length 71;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 DPPCGV 61
Db 1 DPPCGV 6

RESULT 11
US-08-785-065-11
Sequence 11, Application US/08785065
Patent No. 5814451

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goli, Surya K.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/785,065
FILING DATE: Herewith
CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0187 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:

INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 106 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 224
US-08-785-065-11

Query Match 4.8%; Score 6; DB 2; Length 106;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 KRLGVD 56
Db 8 KRLGVD 13

RESULT 12
US-08-906-769-143
Sequence 143, Application US/08906769
Patent No. 6077687

GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
APPLICANT: Silver, Gary
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF
TITLE OF INVENTION: MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 190

CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/906,769
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 143:
SEQUENCE CHARACTERISTICS:
LENGTH: 128 amino acids
TYPE: amino acid
TOPOLOGY: linear

US-08-906-769-143

Query Match 4.8%; Score 6; DB 3; Length 128;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 TKIVFN 20
Db 42 TKIVFN 47

RESULT 13
US-08-906-616-143
Sequence 143, Application US/08906616
Patent No. 6121035

GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
APPLICANT: Silver, Gary
TITLE OF INVENTION: FLEA AMINOPEPTIDASE PROTEINS AND USES THEREOF
NUMBER OF SEQUENCES: 190

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross P.C.
 STREET: 1700 Lincoln Street, Suite 3500
 CITY: Denver
 STATE: Colorado
 COUNTRY: USA
 ZIP: 80203

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/906,616
 FILING DATE: 05-AUG-1997
 CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
 NAME: Connell, Gary J.
 REGISTRATION NUMBER: 32,020
 REFERENCE/DOCKET NUMBER: 2618-25-C2-3

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 143:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 128 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-906-616-143

Query Match 4.8%; Score 6; DB 4; Length 128;

Best Local Similarity 100.0%; Pred. No. 67;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 TKIVFN 20
 Db 42 TKIVFN 47

RESULT 14
 US-08-639-075A-143
 Sequence 143, Application US/08639075A
 ; Patent No. 6150125

GENERAL INFORMATION:

APPLICANT: Grieve, Robert B.
 APPLICANT: Rushlow, Keith E.
 APPLICANT: Wu Hunter, Shirley
 APPLICANT: Frank, Glenn R.
 APPLICANT: Stiegler, Gary
 APPLICANT: Gaines, Patrick J.
 APPLICANT: Silver, Gary
 TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
 NUMBER OF SEQUENCES: 190
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sheridan Ross & McIntosh
 STREET: 1700 Lincoln Street, Suite 3500
 CITY: Denver
 STATE: Colorado
 COUNTRY: USA
 ZIP: 80203

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/012,431
 FILING DATE: 23-Jan-1998
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/639,075

FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Connell, Gary J.

REGISTRATION NUMBER: 32,020
 REFERENCE/DOCKET NUMBER: 2618-25-C2

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 143:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 128 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 143 :
 US-09-012-431-143

REGISTRATION NUMBER: 32,020
 REFERENCE/DOCKET NUMBER: 2618-25-C2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223
 INFORMATION FOR SEQ ID NO: 143:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 128 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-639-075A-143

Query Match 4.8%; Score 6; DB 4; Length 128;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 TKIVFN 20
 Db 42 TKIVFN 47

RESULT 15
 US-09-012-431-143

; Sequence 143, Application US/09012431
 ; Patent No. 6180383
 GENERAL INFORMATION:

APPLICANT: Grieve, Robert B.
 APPLICANT: Rushlow, Keith E.
 APPLICANT: Wu Hunter, Shirley
 APPLICANT: Frank, Glenn R.
 APPLICANT: Stiegler, Gary
 APPLICANT: Gaines, Patrick J.
 APPLICANT: Silver, Gary
 TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
 NUMBER OF SEQUENCES: 190
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sheridan Ross & McIntosh
 STREET: 1700 Lincoln Street, Suite 3500
 CITY: Denver
 STATE: Colorado
 COUNTRY: USA
 ZIP: 80203

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/012,431
 FILING DATE: 23-Jan-1998
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/639,075

FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Connell, Gary J.

REGISTRATION NUMBER: 32,020
 REFERENCE/DOCKET NUMBER: 2618-25-C2

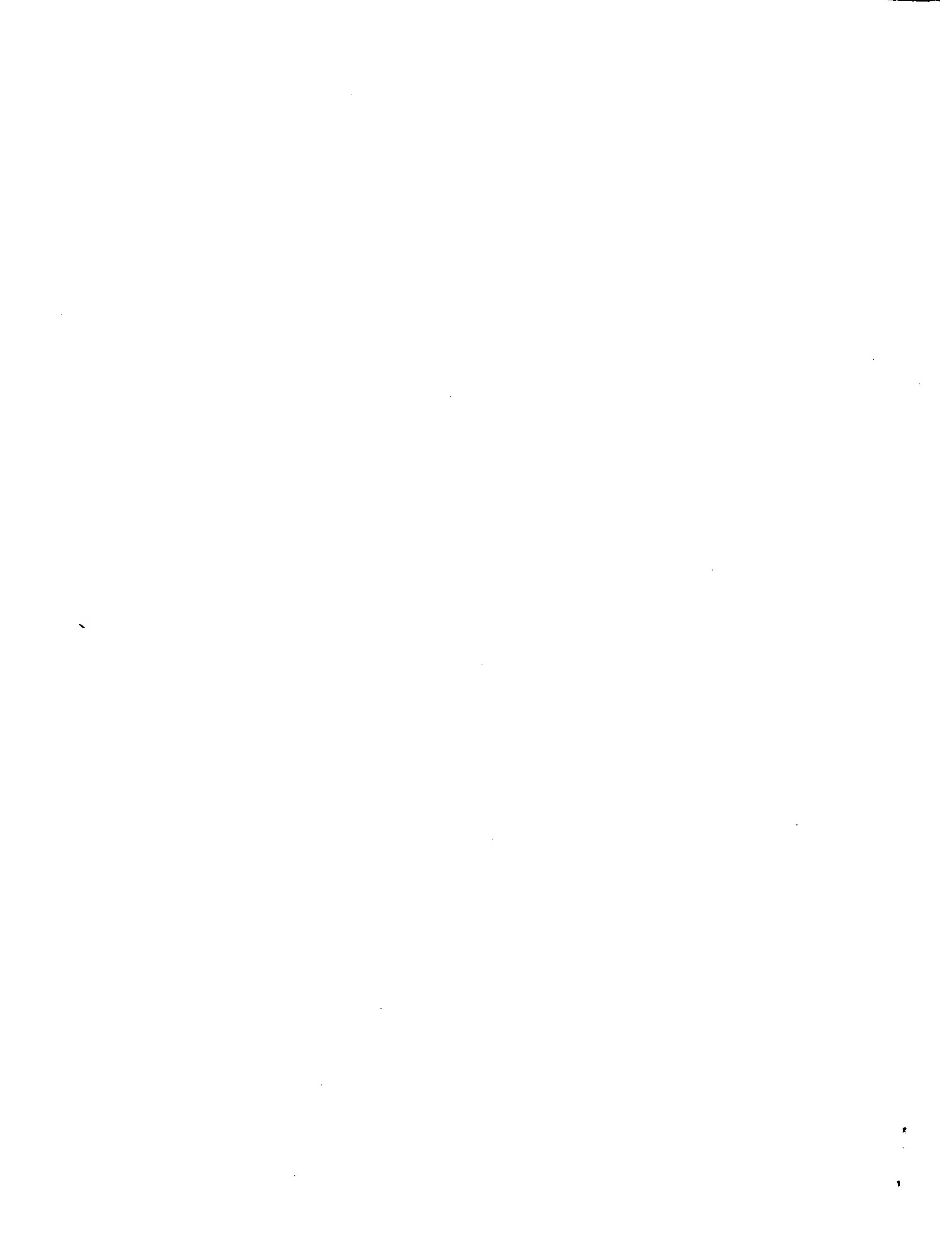
TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 143:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 128 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 143 :

US-09-012-431-143

```
Query Match      4.8%;  Score 6;  DB 4;  Length 128;
Best Local Similarity 100.0%;  Pred. No. 67;
Matches       6;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
               |
QY      15 TKIVFN 20
          |||||
Db      42 TKIVFN 47
```

Search completed: September 24, 2002, 11:03:14
Job time: 158 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2002, 11:01:01 ; Search time 15.96 Seconds
(without alignments)
758.600 Million cell updates/sec

Title: US-09-863-063-2
Perfect score: 126
Sequence: 1 AQSVPPGDIQTQPGTKIVFN.....EWFQGDGMVRRKNLPIEYNP 126

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_71:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	126	100.0	127	T16183	hypothetical prote
2	101	80.2	127	T21640	hypothetical prote
3	101	80.2	127	G88145	protein F58A6.8 [i
4	101	80.2	127	A88165	protein ZK1248.6 [i
5	101	80.2	127	G88686	protein msp-19 [im
6	101	80.2	127	C88688	protein msp-113 [i
7	101	80.2	127	H88688	protein msp-59 [im
8	101	80.2	127	H88689	protein msp-65 [im
9	101	80.2	127	C88689	protein msp-51 [im
10	101	80.2	127	H88792	protein K07F5.1 [i
11	101	80.2	127	H88146	protein C34F11.4 [i
12	101	80.2	127	E88134	protein msp-40 [im
13	101	80.2	127	F88138	protein MSP-31 [im
14	101	80.2	127	D88164	protein msp-142 [i
15	100	79.4	127	A88139	protein ZK546.6 [i
16	93	73.8	190	T16687	major sperm protei
17	90	71.4	127	T24885	hypothetical Prote
18	86	68.3	127	F88801	protein C04G2.4 [i
19	71	56.3	127	C88164	protein K05F1.7 [i
20	68	54.0	127	T16684	major sperm prote
21	59	46.8	127	A88683	protein C09B9.6 [i
22	53	42.1	127	F88146	protein C34F11.6 [i
23	52	41.3	77	F88165	protein ZK1248.4 [i
24	51	40.5	133	T27902	hypothetical prote
25	39	31.0	127	T23486	hypothetical prote
26	23	18.3	127	A45944	major sperm protei
27	23	18.3	127	A45528	major sperm protei
28	23	18.3	127	B45528	hypothetical prote
29	19	15.1	95	T34500	

RESULT 2

Query	Match	Score 100.0%	Length 127;
Qy	1	AQSVPFGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYIKTTNMKRLGVDPGCG	60
Db	2	AQSVPFGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYIKTTNMKRLGVDPGCG	61
Qy	61	VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVIEWNTPDGAARQFRREWFGDGMVRRKNL	120
Db	62	VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVIEWNTPDGAARQFRREWFGDGMVRRKNL	121
Qy	121	PIEYNP	126
Db	122	PIEYNP	127

RESULT 1

Query	Match	Score 100.0%	Length 127;
Qy	1	AQSVPFGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYIKTTNMKRLGVDPGCG	60
Db	2	AQSVPFGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYIKTTNMKRLGVDPGCG	61
Qy	1	VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVIEWNTPDGAARQFRREWFGDGMVRRKNL	120
Db	2	VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVIEWNTPDGAARQFRREWFGDGMVRRKNL	121
Qy	121	PIEYNP	126
Db	122	PIEYNP	127

A; Experimental source: clone F32B6
 R; Swinburne, J.
 submitted to the EMBL Data Library, October 1996
 A; Reference number: Z119947
 A; Accession: T24884
 A; Status: preliminary; translated from GB/EMBL/DDBJ
 A; Residues: 1-127 <STO>
 A; Cross-references: EMBL:Z81122; PIDN:CA03361.1; GSPDB:GN00022; CESP:T13F2.10
 A; Experimental source: clone T13F2
 C; Genetics:
 A; Gene: CESP:F32B6.6; CESP:T13F2.10
 A; Map position: 4
 C; Superfamily: Caenorhabditis elegans major sperm protein

Query Match 80.2%; Score 101; DB 2; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.5e-100;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AQSVPPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 60
 Db 2 AQSVPPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 61
 QY 61 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 101
 Db 62 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 102

RESULT 3
 G88145 protein F58A6.8 [Imported] - Caenorhabditis elegans
 C; Species: Caenorhabditis elegans
 C; Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
 C; Accession: G88145
 R; anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A; Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology
 A; Reference number: A75000; MUID:99069613; PMID:9851916
 A; Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_elegans/
 A; Note: published errata appeared in Science 283, 35, 1999; and
 A; Accession: G88145
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-127 <STO>
 A; Cross-references: GB:chr_II; PIDN:AA96204.1; PID:91255857; GSPDB:GN00020; CESP:F58A6.8
 C; Genetics:
 A; Gene: F58A6.8
 A; Map position: 2
 C; Superfamily: Caenorhabditis elegans major sperm protein

Query Match 80.2%; Score 101; DB 2; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.5e-100;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AQSVPPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 60
 Db 2 AQSVPPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 61
 QY 61 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 101
 Db 62 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 102

RESULT 4
 A88165 protein ZK1248.6 [Imported] - Caenorhabditis elegans
 C; Species: Caenorhabditis elegans
 C; Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
 C; Accession: A88165
 R; anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998

A; Title: Genome sequence of the nematode *C. elegans*: a platform for investigating bio
 R; Swinburne, J.
 submitted to the EMBL Data Library, October 1996
 A; Reference number: Z119947
 A; Accession: T24884
 A; Status: preliminary; translated from GB/EMBL/DDBJ
 A; Residues: 1-127 <STO>
 A; Cross-references: GB:chr_II; PID:9862495; GSPDB:GN00020; CESP:ZK1248.6
 A; Note: similar to *C. elegans* major sperm protein
 C; Genetics:
 A; Gene: ZK1248.6
 A; Map position: 2
 C; Superfamily: Caenorhabditis elegans major sperm protein

Query Match 80.2%; Score 101; DB 2; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.5e-100;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AQSVPPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 60
 Db 2 AQSVPPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 61
 QY 61 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 101
 Db 62 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 102

RESULT 5
 G88686 protein msp-19 [Imported] - Caenorhabditis elegans
 C; Species: Caenorhabditis elegans
 C; Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
 C; Accession: G88686
 R; anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A; Title: Genome sequence of the nematode *C. elegans*: a platform for investigating bio
 A; Reference number: A75000; MUID:99069613; PMID:9851916
 A; Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_elegans/
 A; Note: published errata appeared in Science 283, 35, 1999; and
 A; Accession: G88686
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-127 <STO>
 A; Cross-references: GB:chr_IV; PIDN: AAC26926.1; PID:93329619; GSPDB:GN00022
 C; Genetics:
 A; Gene: msp-19
 A; Map position: 4
 C; Superfamily: Caenorhabditis elegans major sperm protein

Query Match 80.2%; Score 101; DB 2; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.5e-100;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AQSVPPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 60
 Db 2 AQSVPPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 61
 QY 61 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 101
 Db 62 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 102

RESULT 6
 C88688 protein msp-113 [Imported] - Caenorhabditis elegans
 C; Species: Caenorhabditis elegans
 C; Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
 C; Accession: C88688
 R; anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998

A; Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology
 A; Reference number: A75000; MUID:99069613; PMID:9851916
 A; Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_elegans/
 A; Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
 A; Accession: CB8688
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-127 <STO>
 A; Cross-references: GB:chr_IV; PIDN:AAB42255.1; PID:q1825633; GSPDB:GN00022
 C; Genetics:
 A; Gene: msp-59
 A; Map position: 4
 C; Superfamily: Caenorhabditis elegans major sperm protein

Query Match 80.2%; Score 101; DB 2; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.5e-100;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQSVPPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 60
 Db 2 AQSVPPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 61

Qy 61 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVIEWTNTPDGAA 101
 Db 62 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVIEWTNTPDGAA 102

RESULT 9
 C88689 protein msp-51 [imported] - Caenorhabditis elegans
 C; Species: Caenorhabditis elegans
 C; Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
 C; Accession: C88689
 R; anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A; Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology
 A; Reference number: A75000; MUID:99069613; PMID:9851916
 A; Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_elegans/
 A; Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
 A; Accession: H88688
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-127 <STO>
 A; Cross-references: GB:chr_IV; PIDN:AAB42253.1; PID:q1825631; GSPDB:GN00022
 C; Genetics:
 A; Gene: msp-59
 A; Map position: 4
 C; Superfamily: Caenorhabditis elegans major sperm protein

Query Match 80.2%; Score 101; DB 2; Length 127;
 Best Local Similarity 100.0%; Pred. No. 1.5e-100;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQSVPPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 60
 Db 2 AQSVPPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 61

Qy 61 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVIEWTNTPDGAA 101
 Db 62 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVIEWTNTPDGAA 102

RESULT 10
 H88792 protein K07F5.1 [imported] - Caenorhabditis elegans
 C; Species: Caenorhabditis elegans
 C; Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
 C; Accession: H88792
 R; anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A; Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology
 A; Reference number: A75000; MUID:99069613; PMID:9851916

A; Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/projects/C_elegans/ and www_sanger.ac.uk/projects/C_elegans/ and www_sanger.ac.uk/projects/C_elegans/
A; Note: published errata appeared in Science 283, 35, 1999; Science 283, 35, 1999; and
A; Accession: H88146
A; Status: Preliminary
A; Molecule type: DNA
A; Residues: 1-127 <STO>
A; Cross-references: GB:chr_IV; PIDN:CAA94282.1; PID:93878316; GSPDB:GN00022; CESP:K07F5.
C; Genetics:
A; Gene: K07F5.1
A; Map position: 4
C; Superfamily: Caenorhabditis elegans major sperm protein

Query Match Score 101; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.5e-100;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query 1 AQSVPQPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 60
Db 2 AQSVPQPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 61
Db 61 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 101
Db 62 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 102

RESULT 11
H88146
protein C34F11.4 [imported] - Caenorhabditis elegans
C; Species: caenorhabditis elegans
C; Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C; Accession: H88146
R; anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A; Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology
A; Reference number: A75000; MUID:99069613; PMID:9851916
A; Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/projects/C_elegans/
A; Note: published errata appeared in Science 283, 35, 1999; Science 283, 35, 1999; and
A; Accession: H88146
A; Status: Preliminary
A; Molecule type: DNA
A; Residues: 1-127 <STO>
A; Cross-references: GB:chr_II; PIDN:AA85761.1; PID:g1166627; GSPDB:GN00020; CESP:C34F11.
C; Genetics:
A; Gene: C34F11.4
A; Map position: 2
C; Superfamily: Caenorhabditis elegans major sperm protein

Query Match Score 101; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.5e-100;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query 1 AQSVPQPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 60
Db 2 AQSVPQPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 61
Db 61 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 101
Db 62 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 102

RESULT 12
E88134
protein msp-40 [imported] - Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C; Accession: E88134
R; anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A; Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology
A; Reference number: A75000; MUID:99069613; PMID:9851916

A; Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/projects/C_elegans/
A; Note: published errata appeared in Science 283, 35, 1999; Science 283, 35, 1999; and
A; Accession: E88134
A; Status: Preliminary
A; Molecule type: DNA
A; Residues: 1-127 <STO>
A; Cross-references: GB:chr_II; PIDN:AA93398.1; PID:g1203940; GSPDB:GN00020; CESP:C33F
C; Genetics:
A; Gene: msp-40
A; Map position: 2
C; Superfamily: Caenorhabditis elegans major sperm protein

Query Match Score 101; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.5e-100;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query 1 AQSVPQPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 60
Db 2 AQSVPQPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 61
Db 61 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 101
Db 62 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 102

RESULT 13
F88138
protein MSP-31 [imported] - Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C; Accession: F88138
R; anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A; Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology
A; Reference number: A75000; MUID:99069613; PMID:9851916
A; Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/projects/C_elegans/
A; Note: published errata appeared in Science 283, 35, 1999; Science 283, 35, 1999; and
A; Accession: H88138
A; Status: Preliminary
A; Molecule type: DNA
A; Residues: 1-127 <STO>
A; Cross-references: GB:chr_II; PIDN:AA83175.1; PID:g1109821; GSPDB:GN00020; CESP:RO5F
C; Genetics:
A; Gene: MSP-31
A; Map position: 2
C; Superfamily: Caenorhabditis elegans major sperm protein

Query Match Score 101; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.5e-100;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query 1 AQSVPQPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 60
Db 2 AQSVPQPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYGIKTTNMKRLGVDPGCG 61
Db 61 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 101
Db 62 VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 102

RESULT 14
D88164
protein msp-142 [imported] - Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C; Accession: D88164
R; anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A; Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology
A; Reference number: A75000; MUID:99069613; PMID:9851916

A; Reference number: A75000; MUID:9069613; PMID:9851916
 A; Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_ele
 A; Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A; Accession: D88164
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-127 <STO>
 A; Cross-references: GB:chr_II; PID:9868174; GSPDB:GN00020; CESP:K05FL.2
 A; Note: K05FL.2
 C; Genetics:
 A; Gene: msp-142
 A; Map position: 2
 C; Superfamily: Caenorhabditis elegans major sperm protein

```
Query Match 80.2%; Score 101; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.5e-100;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1  AQSVPPGDIQTQPGTKIVFNAPYDDKHTYHIVINSSARRIGYGIKTTNMKRLGVDPGCG 60
Db  2  AQSVPPGDIQTQPGTKIVFNAPYDDKHTYHIVINSSARRIGYGIKTTNMKRLGVDPGCG 61

Qy  61  VLDPEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 101
Db  62  VLDPEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 102
```

RESULT 15
 A88139
 protein ZK546.6 [imported] - Caenorhabditis elegans
 C; Species: Caenorhabditis elegans
 C; Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
 C; Accession: A88139
 R; anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A; Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology
 A; Reference number: A75000; MUID:9069613; PMID:9851916
 A; Note: see websites genome.wustl.edu/gsc/C_elegans/ and www_sanger.ac.uk/Projects/C_ele
 A; Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A; Accession: A88139
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-127 <STO>
 A; Cross-references: GB:chr_II; PID:g868205; GSPDB:GN00020; CESP:ZK546.6
 A; Note: Major sperm protein
 C; Genetics:
 A; Gene: ZK546.6
 A; Map position: 2
 C; Superfamily: Caenorhabditis elegans major sperm protein

```
Query Match 79.4%; Score 100; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.8e-99;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2  QSVPPGDIQTQPGTKIVFNAPYDDKHTYHIVINSSARRIGYGIKTTNMKRLGVDPGCG 61
Db  3  QSVPPGDIQTQPGTKIVFNAPYDDKHTYHIVINSSARRIGYGIKTTNMKRLGVDPGCG 62

Qy  62  LDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 101
Db  63  LDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAA 102
```

Search completed: September 24, 2002, 11:03:36
 Job time: 155 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: September 24, 2002, 11:02:21 ; Search time 11.85 Seconds
(without alignments)
411.702 Million cell updates/sec

Title: US-09-863-063-2
Perfect score: 126
Sequence: 1AQSVPPGDIQTQPGTKIVFN.....EWFQGDGMVRRKNNLPIEYNP 126

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	101	80.2	126	1	MS31_CAEEL	P53017 caenorhabdi
2	93	73.8	190	1	MS32_CAEEL	P53018 caenorhabdi
3	68	54.0	126	1	MS33_CAEEL	P53019 caenorhabdi
4	56	44.4	126	1	MS30_CAEEL	P05634 caenorhabdi
5	53	42.1	126	1	MS56_CAEEL	P05635 caenorhabdi
6	39	31.0	126	1	MS38_CAEEL	P53020 caenorhabdi
7	23	18.3	126	1	MSP1_ASCSU	P27439 ascaris suu
8	23	18.3	126	1	MSP1_ONCVO	P13262 onchocerca
9	23	18.3	126	1	MSP2_ASCSU	P27440 ascaris suu
10	23	18.3	126	1	MSP2_ONCVO	P13263 onchocerca
11	19	15.1	125	1	MSP1_GLORO	P53021 globodera r
12	19	15.1	125	1	MSP2_GLORO	P53022 globodera r
13	19	15.1	125	1	MSP3_GLORO	P53023 globodera r
14	8	6.3	265	1	KNH1_CANGA	074684 candida gla
15	8	6.3	454	1	GABA_HUMAN	Q06546 homo sapien
16	8	6.3	454	1	GABA_MOUSE	Q00422 mus musculu
17	7	5.6	178	1	R20K_CLOPA	P23161 clostridium
18	7	5.6	249	1	COBS_MYCTU	Q10397 mycobacteri
19	7	5.6	511	1	NEK3_MOUSE	Q02379 bos taurus
20	7	5.6	1066	1	T145_YEAST	P46677 saccharomyces
21	7	5.6	1106	1	ACLY_CAEEL	P53585 caenorhabdi
22	6	4.8	64	1	RLL1_STR1A	087734 streptomyce
23	6	4.8	102	1	HCC2_CRYCO	Q01238 cryptheco
24	6	4.8	105	1	NIPM_BOVIN	Q02379 bos taurus
25	6	4.8	115	1	MOTI_BOVIN	062820 bos taurus
26	6	4.8	121	1	KDGL_ECOLI	P00556 escherichia
27	6	4.8	129	1	VMOR_BPMU	P23848 bacterioph
28	6	4.8	135	1	Y360_METJU	Q57806 methanococcus
29	6	4.8	137	1	ATPE_PICAB	Q47036 picea abies
30	6	4.8	144	1	MARR_SALTY	Q56069 salmonella
31	6	4.8	150	1	SPOA_BACCE	P52930 bacillus ce
32	6	4.8	177	1	GPIX_HUMAN	P14770 homo sapien
33	6	4.8	182	1	RRF_SYN3	P74456 synechocyst

ALIGNMENTS

RESULT ID	MS31_CAEEL	STANDARD;	PRT;	126 AA.
AC	P53017;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Major sperm protein 31/40/142 (MSP).			
GN	(MSP-31 OR R05F9.13) AND (MSP-40 OR C33F10.9) AND (MSP-51 OR ZK354.5)			
GN	AND (MSP-59 OR ZK354.11) AND (MSP-65 OR ZK354.1) AND (MSP-113 OR ZK354.4) AND (MSP-142 OR K05F1.2) AND C34F11.4 AND F58A6.8 AND K07F5.1			
GN	AND ZK1248.6.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;			
OX	OCBL_TAXID=6239;			
RA	Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.			
RP	[11]			
RC	SEQUENCE FROM N.A. STRAIN=BRISTOL N2;			
RL	Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.			
RN	[12]			
RP	SEQUENCE FROM N.A. STRAIN=BRISTOL N2;			
RC	Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.			
RA	Wohldmann P.; Waterston R.;			
RL	[13]			
RN	SEQUENCE FROM N.A. STRAIN=BRISTOL N2;			
RP	Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.			
RC	[4]			
RA	SEQUENCE FROM N.A. STRAIN=BRISTOL N2;			
RL	Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A. STRAIN=BRISTOL N2;			
RC	Bentley D.; Waterston R.;			
RA	Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RL	[6]			
RN	SEQUENCE FROM N.A. STRAIN=BRISTOL N2;			
RP	Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.			
RC	[7]			
RA	Latreille P.; Hembry C.;			
RL	Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.			
RN	[8]			
RP	SEQUENCE FROM N.A. STRAIN=BRISTOL N2;			
RC	Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.			
RA	[9]			
RL	SEQUENCE FROM N.A. STRAIN=BRISTOL N2;			
RN	[10]			
RP	SEQUENCE FROM N.A. STRAIN=BRISTOL N2;			
RC	Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.			
RA	[11]			

CC -!- FUNCTION: CENTRAL COMPONENT IN MOLECULAR INTERACTIONS UNDERLYING
CC SPERM CRAWLING. FORMS AN EXTENSIVE FILAMENT SYSTEM THAT EXTENDS
CC FROM SPERM VILLIPODA, ALONG THE LEADING EDGE OF THE PSEUDOPOD.
CC -!- TISSUE SPECIFICITY: SPERM.
CC -!- SIMILARITY: BELOWS TO THE NEMATODE MSP FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; U29244; AAC71087.1; -. DR WormPep; C34F11.4; CE02806.
DR EMBL; U29377; AAA68711.1; -. DR WormPep; F58A6.8; CE02806.
DR EMBL; U46753; AAA85761.1; -. DR WormPep; K05F1.2; CE02806.
DR EMBL; U49830; AAK31477.1; -. DR WormPep; K07F5.1; CE02806.
DR EMBL; U53339; AAA96204.1; -. DR WormPep; R05F9.13; CE02806.
DR EMBL; U41533; AAA83175.1; -. DR WormPep; U88172; AAB42253.1; -. DR InterPro; IPR00535; MSP_domain.
DR EMBL; Z70284; CAA94282.1; -. DR WormPep; U88172; AAB42254.1; -. DR Pfam; PF00635; MSP_domain; 1.
DR EMBL; U88172; AAB42255.1; -. DR Cytoskeleton; Sperm; Multigene family.
DR EMBL; U88172; AAB42256.1; -. DR Sequence 190 AA; 939C5BBED79C54BA CRC64;
DR HSSP; P27439; 1MSP.
DR WormPep; C33F10.9; 3MSP.
DR WormPep; C34F11.4; CE02806.
DR WormPep; F58A6.8; CE02806.
DR WormPep; K05F1.2; CE02806.
DR WormPep; K07F5.1; CE02806.
DR WormPep; R05F9.13; CE02806.
DR WormPep; ZK1248.6; CE02806.
DR WormPep; 2K354.1; CE09978.
DR WormPep; 2K354.4; CE09978.
DR WormPep; ZK354.5; CE09978.
DR WormPep; ZK354.11; CE09978.
DR InterPro; IPR00535; MSP_domain.
DR Pfam; PF00635; MSP_domain; 1.
KW Cytoskeleton; Acetylation; Sperm; Multigene family.
FT INIT_MET 0 0 BY SIMILARITY.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
SQ SEQUENCE 126 AA; 14078 MW; QF069631D8559AB7 CRC64;

Query Match 73.8%; Score 93; DB 1; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.7e-92;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DR InterPro; IPR00535; MSP_domain.
DR Pfam; PF00635; MSP_domain; 1.
DR Cytoskeleton; Sperm; Multigene family.
SQ Sequence 190 AA; 939C5BBED79C54BA CRC64;

RESULT 3
MS33_CAEEL
ID MS33_CAEEL STANDARD PRT; 126 AA.
AC P53019;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major sperm protein 33 (MSP).
GN MSP-33 OR R05F9.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderaidae; Caenorhabditis.
NCBI_TaxID=6239;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=BRISTOL N2;
RA Hallsworth K.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: CENTRAL COMPONENT IN MOLECULAR INTERACTIONS UNDERLYING
CC SPERM CRAWLING. FORMS AN EXTENSIVE FILAMENT SYSTEM THAT EXTENDS
CC FROM SPERM VILLIPODA, ALONG THE LEADING EDGE OF THE PSEUDOPOD.
CC -!- TISSUE SPECIFICITY: SPERM.
CC -!- SIMILARITY: BELOWS TO THE NEMATODE MSP FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

RESULT 2
MS32_CAEEL
ID MS32_CAEEL STANDARD PRT; 190 AA.
AC P53018;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major sperm protein 32 (MSP).
GN MSP-32 OR R05F9.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderaidae; Caenorhabditis.
NCBI_TaxID=6239;
[1]

CC	EMBL; U41533; AAA83165.1;	-.	Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DR	HSSP; P27439; 3MSP.		
DR	WormPep; R05F9.8; CE04811.		
DR	InterPro; IPR000535; MSP_domain.		
DR	Pfam; PF00635; MSP_domain; 1.		
KW	Cytoskeleton; Acetylation; Sperm; Multigene family.		
FT	INIT_MET 0 0 BY SIMILARITY.		
MOD_RES	1 1 ACETYLATION (BY SIMILARITY).		
SQ	SEQUENCE 126 AA; 14112 MW; 0F0C9C31D25F9AB7 CRC64;		
Query Match 54.0%; Score 68; DB 1; Length 126;			
Best Local Similarity 100.0%; Pred. No. 8.9e 66;			
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
YQ	1 AQSVPPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYIKTTNMKRLGVDPGCG 60		
Db	1 AQSVPPGDIQTQPGTKIVFNAPYDDKHTYHIKVINSSARRIGYIKTTNMKRLGVDPGCG 60		
YQ	61 VLDPKEAV 68		
Db	61 VLDPKEAV 68		
RESULT 4	MS10_CAEEL STANDARD; PRT; 126 AA.		
ID	MS10_CAEEL		
AC	P05634;		
DT	01-NOV-1988 (Rel. 09, Created)		
DT	01-AUG-1992 (Rel. 23, Last sequence update)		
DT	01-OCT-1996 (Rel. 34, Last annotation update)		
DE	Major sperm protein 10 (MSP).		
GN	MSP-10		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae; Rhabditidae; Pelederinae; Caenorhabditis.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BRISTOL N2;		
RX	PubMed=6325882;		
RA	Klass M.R., Kinsley S., Lopez L.C.;		
RT	"Isolation and characterization of a sperm-specific gene family in the nematode Caenorhabditis elegans."		
RL	Mol. Cell. Biol. 4:529-537(1984).		
CC	-!- FUNCTION: CENTRAL COMPONENT IN MOLECULAR INTERACTIONS UNDERLYING SPERM CRAWLING. FORMS AN EXTENSIVE FILAMENT SYSTEM THAT EXTENDS FROM SPERM VILLIPODA, ALONG THE LEADING EDGE OF THE PSEUDOPOD.		
CC	-!- TISSUE SPECIFICITY: SPERM.		
CC	-!- MISCELLANEOUS: AROUND 30 MSP ISOFORMS MAY EXIST IN C. ELEGANS.		
CC	-!- SIMILARITY: BELONGS TO THE NEMATODE MSP FAMILY.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	CC DR EMBL; K02618; AAA28116.1; -.		
CC	CC DR HSSP; P27439; 3MSP.		
CC	CC DR InterPro; IPR000535; MSP_domain.		
CC	CC DR Pfam; PF00635; MSP_domain; 1.		
CC	CC KW Cytoskeleton; Acetylation; Sperm; Multigene family.		
CC	CC FT INIT_MET 0 0 BY SIMILARITY.		
CC	CC FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).		
CC	CC SO SEQUENCE 126 AA; 14063 MW; 82F7B36A4D80C5C8 CRC64;		
CC	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	CC DR EMBL; K02617; AAA28115.1; -.		
CC	CC DR HSSP; P27439; 3MSP.		
CC	CC DR InterPro; IPR000535; MSP_domain.		
CC	CC DR Pfam; PF00635; MSP_domain; 1.		
KW	Cytoskeleton; Acetylation; Sperm; Multigene family.		
FT	INIT_MET 0 0 BY SIMILARITY.		
FT	MOD_RES 1 1 ACETYLATION (BY SIMILARITY).		
SQ	SEQUENCE 126 AA; 14093 MW; 835CA9DBFD80D5D9 CRC64;		
Query Match 44.4%; Score 56; DB 1; Length 126;			
Best Local Similarity 100.0%; Pred. No. 7.1e-53;			
YQ	46 KTtNmKRLGVDPGCVLDPKEAVLLAVSCDAFAFGQQEDTNNDRITVIEWTNTPDGAA 101		
Db	46 KTtNmKRLGVDPGCVLDPKEAVLLAVSCDAFAFGQQEDTNNDRITVIEWTNTPDGAA 101		
RESULT 5	MS56_CAEEL STANDARD; PRT; 126 AA.		
ID	MS56_CAEEL		
AC	P05635;		
DT	01-NOV-1988 (Rel. 09, Created)		
DT	01-AUG-1992 (Rel. 23, Last sequence update)		
DE	Major sperm protein 56 (MSP).		
GN	MSP-56.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae; Rhabditidae; Pelederinae; Caenorhabditis.		
NCBI_TAXID=6239;			
RN [1]	SEQUENCE FROM N.A.		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BRISTOL N2;		
RX	PubMed=6325882;		
RA	Klass M.R., Kinsley S., Lopez L.C.;		
RT	"Isolation and characterization of a sperm-specific gene family in the nematode Caenorhabditis elegans."		
RL	Mol. Cell. Biol. 4:529-537(1984).		
CC	-!- FUNCTION: CENTRAL COMPONENT IN MOLECULAR INTERACTIONS UNDERLYING SPERM CRAWLING. FORMS AN EXTENSIVE FILAMENT SYSTEM THAT EXTENDS FROM SPERM VILLIPODA, ALONG THE LEADING EDGE OF THE PSEUDOPOD.		
CC	-!- TISSUE SPECIFICITY: SPERM.		
CC	-!- MISCELLANEOUS: AROUND 30 MSP ISOFORMS MAY EXIST IN C. ELEGANS.		
CC	-!- SIMILARITY: BELONGS TO THE NEMATODE MSP FAMILY.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	DR EMBL; K02617; AAA28115.1; -.		
CC	DR HSSP; P27439; 3MSP.		
CC	DR InterPro; IPR000535; MSP_domain.		
CC	DR Pfam; PF00635; MSP_domain; 1.		
KW	Cytoskeleton; Acetylation; Sperm; Multigene family.		
FT	INIT_MET 0 0 BY SIMILARITY.		
FT	MOD_RES 1 1 ACETYLATION (BY SIMILARITY).		
SQ	SEQUENCE 126 AA; 14093 MW; 82F7B36A4D80C5C8 CRC64;		
Query Match 44.4%; Score 56; DB 1; Length 126;			
Best Local Similarity 100.0%; Pred. No. 7.1e-53;			
YQ	49 NMKRGLGVDPGCVLDPKEAVLLAVSCDAFAFGQQEDTNNDRITVIEWTNTPDGAA 101		
Db	49 NMKRGLGVDPGCVLDPKEAVLLAVSCDAFAFGQQEDTNNDRITVIEWTNTPDGAA 101		
RESULT 6	MS38_CAEEL STANDARD; PRT; 126 AA.		
ID	MS38_CAEEL		
AC	P53020;		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DE	Major sperm protein 38 (MSP).		
GN	MSP-38 OR K08F4.8.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;		

RA Scott A.L., Dinman J., Sussman D.J., Yenbutr P., Ward S.;
 RT "Major sperm protein genes from Onchocerca volvulus.";
 RL Mol. Biochem. Parasitol. 36:119-126(1989).
 -!- FUNCTION: CENTRAL COMPONENT IN MOLECULAR INTERACTIONS UNDERLYING
 SPERM CRAWLING. FORMS AN EXTENSIVE FILAMENT SYSTEM THAT EXTENDS
 FROM SPERM VILLIPODA, ALONG THE LEADING EDGE OF THE PSEUDOPOD.
 CC -!- TISSUE SPECIFICITY: SPERM.
 CC -!- SIMILARITY: BELONGS TO THE NEMATODE MSP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR J04662; -; NOT_ANNOTATED_CDS.
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D308E525B511EC81 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D308E525B511EC81 CRC64;
 SQ -----
 Query Match 18.3%; Score 23; DB 1; Length 126;
 Best Local Similarity 100.0%; Pred. No. 2.1e-17;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 103 QFRREWFQGDGMVRRKNLPIEYN 125
 Db 103 QFRREWFQGDGMVRRKNLPIEYN 125
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D308E525B511EC81 CRC64;
 SQ -----
 RESULT 10
 MSP2_ONCVO STANDARD PRT; 126 AA.
 ID MSP2_ONCVO STANDARD PRT; 126 AA.
 AC P13263;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-AUG-1992 (Rel. 23, Last annotation update)
 DE Major sperm protein 2 (MSP2).
 OS Onchocerca volvulus.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Onchocerca.
 OC NCBITaxID=6282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89365002; PubMed=2770787;
 RA Scott A.L., Dinman J., Sussman D.J., Yenbutr P., Ward S.;
 RT "Major sperm protein genes from Onchocerca volvulus."
 RL Mol. Biochem. Parasitol. 36:119-126(1989).
 -!- FUNCTION: CENTRAL COMPONENT IN MOLECULAR INTERACTIONS UNDERLYING
 SPERM CRAWLING. FORMS AN EXTENSIVE FILAMENT SYSTEM THAT EXTENDS
 FROM SPERM VILLIPODA, ALONG THE LEADING EDGE OF THE PSEUDOPOD.
 CC -!- TISSUE SPECIFICITY: SPERM.
 CC -!- SIMILARITY: BELONGS TO THE NEMATODE MSP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J04663; AA29421.1; -.
 DR PIR; B45528; B45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; B45528; B45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 Query Match 18.3%; Score 23; DB 1; Length 126;
 Best Local Similarity 100.0%; Pred. No. 2.1e-17;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 103 QFRREWFQGDGMVRRKNLPIEYN 125
 Db 103 QFRREWFQGDGMVRRKNLPIEYN 125
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 RESULT 11
 MSP1_GLORO STANDARD PRT; 125 AA.
 ID MSP1_GLORO STANDARD PRT; 125 AA.
 AC P53021;
 DR -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR000535; MSP_domain.
 DR pfam; PF00635; MSP_domain; 1.
 KW Cytoskeleton; Acetylation; Sperm; Multigene family.
 INIT_MET 0 BY SIMILARITY.
 MOD_RES 1 ACETYLAION (BY SIMILARITY).
 FT SEQUENCE 126 AA; 14211 MW; D40F9552B561E9D4 CRC64;
 SQ -----
 DR PIR; A45528; A45528.<br

Query Match 15.1%; Score 19; DB 1; Length 125;
 Best Local Similarity 100.0%; Pred. No. 4.1e-13;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 EWFQGDGMVRKNLPIEYN 125
 Db 106 EWFQGDGMVRKNLPIEYN 124

RESULT 14
 KNH1_CANGA STANDARD; PRT; 265 AA.

ID KNH1_CANGA
 AC O74684;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Cell wall synthesis protein KNH1 precursor.
 GN KNH1.
 OS Candida glabrata (Yeast) (Torulopsis glabrata);
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5478;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 2001;
 RX MEDLINE=98422452; PubMed=9748432;
 RA Nagahashi S., Lussier M., Bussey H.;
 RT "Isolation of Candida glabrata homologs of the Saccharomyces cerevisiae KRE9 and KNH1 genes and their involvement in cell wall beta-1,6-glucan synthesis";
 RT J. Bacteriol. 180:5020-5029(1998).
 CC -!- FUNCTION: INVOLVED IN CELL WALL BETA(1->6) GLUCAN SYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: SECRETED, PROBABLY FOUND AT CELL SURFACE (BY SIMILARITY).
 CC -!- PTM: O-GLYCOSYLATED (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE KRE9 / KNH1 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: BELONGS TO THE KRE9 / KNH1 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AF064252; AAC64009.1; -.
 DR Glycoprotein; Cell wall; Signal.
 FT SIGNAL 1 14 POTENTIAL.
 FT CHAIN 15 265 CELL WALL SYNTHESIS PROTEIN KNH1.
 SQ SEQUENCE 265 AA; 29046 MW; 6A13BCD6BC0AE2BC CRC64;

Query Match 6.3%; Score 8; DB 1; Length 265;
 Best Local Similarity 100.0%; Pred. No. 0.52%;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 QTPGPTKI 17
 Db 175 QTPGPTKI 182

RESULT 15
 GABA_HUMAN STANDARD; PRT; 454 AA.

ID GABA_HUMAN
 AC Q06546; Q12939;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE GA binding protein alpha chain (GABP-alpha subunit) (Transcription factor E4TF1-60) (Nuclear respiratory factor-2 subunit alpha).
 GN E4TF1A OR GABPA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93180783; PubMed=8441384;
 RA Watanabe H., Sawada J.-I., Yano K.-I., Yamauchi K., Goto M.,
 RA Handa H.;
 RT "cDNA cloning of transcription factor E4TF1 subunits with Ets and notch motifs.";
 RT Mol. Cell. Biol. 13:1385-1391(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95097980; PubMed=7799916;
 RA Gugneja S., Virbasius J.V., Scarpaulla R.C.;
 RT "Four structurally distinct, non-DNA-binding subunits of human nuclear respiratory factor 2 share a conserved transcriptional activation domain.";
 RT Mol. Cell. Biol. 15:102-111(1995).
 CC -!- FUNCTION: TRANSCRIPTION FACTOR CAPABLE OF INTERACTING WITH PURINE RICH REPEATS (GA REPEATS). NECESSARY FOR THE EXPRESSION OF THE ADENOVIRUS E4 GENE.
 CC -!- SUBUNIT: HETEROCTETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: CONTAINS 1 POINTED (PNT) DOMAIN.
 CC -!- SIMILARITY: BELONGS TO THE ETS FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 POINTED (PNT) DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; D13318; BAA02575.1; -.
 DR EMBL; U13044; AAA65706.1; -.
 DR HSSP; Q00422; IAWC.
 DR TRANSFAC; T01390; -.
 DR MIM; 600609; -.
 DR InterPro; IPR000418; Ets.
 DR InterPro; IPR002341; HSF_ETS.
 DR InterPro; IPR003118; SAM_PNT.
 DR Pfam; PF00178; Ets; 1.
 DR Pfam; PF02198; SAM_PNT; 1.
 DR PRINTS; PR00454; ETSDOMAIN.
 DR SMART; SMD00413; Ets; 1.
 DR SMART; SMD0251; SAM_PNT; 1.
 DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
 DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
 DR PROSITE; PS50061; ETS_DOMAIN_3; 1.
 KW Transcription regulation; DNA-binding; Nuclear protein.
 FT DOMAIN 170 251 POINTED.
 FT DNA_BIND 320 400 ETS-DOMAIN.
 FT CONFLICT 289 290 SS -> RC (IN REF. 2).
 FT CONFLICT 440 440 A -> V (IN REF. 2).
 SQ SEQUENCE 454 AA; 51295 MW; 1AF2ABBBBC79191DD CRC64;

Query Match 6.3%; Score 8; DB 1; Length 454;
 Best Local Similarity 100.0%; Pred. No. 0.82%;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 31 IKVINS SA 38
 Db 284 IKVINS SA 291

Search completed: September 24, 2002, 11:04:26
 Job time: 125 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2002, 11:01:31 ; Search time 24.47 Seconds
(without alignments)
890.779 Million cell updates/sec

Title: US-09-863-063-2

Perfect score: 126
Sequence: 1 AQSVPGGDIQTPGTRIVFN.....ENFQGDGMVRRKNLPIEYNP 126

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 5622222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5622222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL_19;*

- 1: sp_archaea;*
- 2: sp_bacteria;*
- 3: sp_fungi;*
- 4: sp_human;*
- 5: sp_invertebrate;*
- 6: sp_mammal;*
- 7: sp_mhc;*
- 8: sp_organelle;*
- 9: sp_phage;*
- 10: sp_plant;*
- 11: sp_rabbit;*
- 12: sp_virus;*
- 13: sp_vertbrate;*
- 14: sp_unclassified;*
- 15: sp_rviruse;*
- 16: sp_bacteriap;*
- 17: sp_archeap;*

ALIGNMENTS

RESULT 1

ID	Q19832	PRELIMINARY;	PRT;	127 AA.
AC	Q19832;			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)			
DE	HYPOTHETICAL 14.2 KDA PROTEIN.			
GN	F26G1.7.			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;			
OC	Rhabditidae; Peloderaiae; Caenorhabditis.			
NCBI_TaxID	6239;			
[1]				
RN		SEQUENCE FROM N.A.		
RP		STRAIN=BRISTOL N2;		
RC		MEDLINE=99069613; PubMed=9851916;		
RX				
RA		None;		
RT		"Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium." Science 282:2012-2018(1998).		
RL		[3]		
[2]				
RN		SEQUENCE FROM N.A.		
RP		STRAIN=BRISTOL N2;		
RC		Chissoe S.;		
RA		"The sequence of C. elegans cosmid F26G1.";		
RT		Submitted (APR-1995) to the EMBL/GenBank/DDJB databases.		
RL				
RN				

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match	Length	DB	ID	Description
1	126	100.0	127	5	Q19832	Q19832 caenorhabdi
2	101	80.2	127	5	Q9TVW5	Q9TVW5 caenorhabdi
3	101	80.2	130	5	Q9N2M1	Q9N2M1 caenorhabdi
4	100	79.4	127	5	Q23519	Q23519 caenorhabdi
5	90	71.4	127	5	Q94053	Q94053 caenorhabdi
6	86	68.3	127	5	Q27280	Q27280 caenorhabdi
7	71	56.3	127	5	Q21244	Q21244 caenorhabdi
8	59	46.8	127	5	Q17856	Q17856 caenorhabdi
9	53	42.1	127	5	Q18461	Q18461 caenorhabdi
10	52	41.3	77	5	Q95PJ7	Q95PJ7 caenorhabdi
11	52	41.3	127	5	Q95XU7	Q95XU7 caenorhabdi
12	51	40.5	133	5	Q23524	Q23524 caenorhabdi
13	22	17.5	107	5	Q9GNW0	Q9GNW0 onchocerca
14	22	17.5	107	5	Q9GNV9	Q9GNV9 onchocerca
15	22	17.5	107	5	Q23428	Q23428 caenorhabdi
16	19	15.1	95	5		

Query Match	1	AQSVPPIQDIQTQPGTKIVFNAPYDDKHTYHIKVINSARRIGYGIKTTNMKRLGVDPGCG	Score 126; DB 5; Length 127;
Best Local Similarity	100.0%	Pred. No. 6.1e-128;	
Matches	126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Db	2	AQSVPPIQDIQTQPGTKIVFNAPYDDKHTYHIKVINSARRIGYGIKTTNMKRLGVDPGCG	60
QY	61	VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAARQFREWFQGDGMVRKNL	120
Db	62	VLDPKEAVLLAVSCDAFAFGQEDTNNDRITVEWTNTPDGAARQFREWFQGDGMVRKNL	121
QY	121	PIEYNP 126	
Db	122	PIEYNP 127	
RESULT	2	Q9TVW5	PRELIMINARY; PRT; 127 AA.
ID	Q9TVW5		
AC	Q9TVW5;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
RR	T13F2.10 PROTEIN		
DR	T13F2.10 OR F32B6.6.		
GN	T13F2.10		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;		
OC	Rhabditidae; Pelerodinae; Caenorhabditis.		
NCBI_TAXID	6239;		
[1]			
SEQUENCE FROM N.A.			
Swinburne J.;			
Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.			
[2]			
SEQUENCE FROM N.A.			
MEMLINE=9069613; PubMed=9851916;			
none;			
"Genome sequence of the nematode <i>C.elegans</i> : A platform for investigating biology.";			
Science 282:2012-2018(1998).			
[3]			
SEQUENCE FROM N.A.			
Basham V.M.;			
Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.			
EMBL; Z81122; CAB03361.1;			
EMBL; Z81074; CAB03037.1;			
HSSP; P27439; 3MSP.			
InterPro; IPR000535; MSP_domain.			
PFam; PF00635; MSP_domain; 1.			
DR	SEQUENCE 127 AA; 14181 MW;	69F6962335A5B147 CRC64;	
RR			
RESULT	4	Q23519	PRELIMINARY; PRT; 127 AA.
ID	Q23519		
AC	Q23519;		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	HYPOTHETICAL_14.2 KDA PROTEIN.		
GN	ZK546.6.		
OS	Caenorhabditis elegans.		
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;		
OC	Rhabditidae; Pelerodinae; Caenorhabditis.		
NCBI_TAXID	6239;		
[1]			
SEQUENCE FROM N.A.			
Strain-Bristol N2;			
Medline=9069613; PubMed=9851916;			
None;			
"Genome sequence of the nematode <i>C.elegans</i> : A platform for investigating biology.";			
Science 282:2012-2018(1998).			
[2]			
SEQUENCE FROM N.A.			
Strain-Bristol N2;			
Hallsworth K.;			
"The sequence of <i>C.elegans</i> cosmid ZK546.";			
Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.			
[3]			
SEQUENCE FROM N.A.			
Strain-Bristol N2;			
Waterston R.;			
RR	RESULT 3		
ID	Q9N2M1	PRELIMINARY;	PRT; 130 AA.
AC	Q9N2M1;		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		

"Direct Submission.";			
Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.			
EMBL; U29380; AAA68739.1; -.			
HSSP; P27439; 3MSP.			
InterPro; IPR000535; MSP_domain.			
Pfam; PF00635; MSP_domain; 1.			
Hypothetical protein.			
SEQUENCE 127 AA; 14239 MW; 69F6962DCEACCI52 CRC64;			
Query Match 79.4%; Score 100; DB 5; Length 127;			
Best Local Similarity 100.0%; Pred. No. 7e-100;			
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
[1] SEQUENCE FROM N.A.			
RA Hembry C.;			
RP Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.			
RN [2]			
RN SEQUENCE FROM N.A.			
RN RA RX MEDLINE=94150718; PubMed=7906398;			
RA RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,			
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,			
RA RA Craxton M., Dear S., Du Z., Durbin R., Favelllo A., Fulton L.,			
RA RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,			
RA RA Jones M., Kirsten J., Laister N., Latreille P.,			
RA RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,			
RA RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,			
RA RA Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,			
RA RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,			
RA RA Watson A., Weinstock L., Wilkison-Sproat J., Wohldman P.;			
RT RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans."			
RL RL Nature 368:32-38(1994).			
RN [3]			
RN SEQUENCE FROM N.A.			
RN RA McMurray A.;			
RN RL Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.			
RN [4]			
RN SEQUENCE FROM N.A.			
RA Hembry C.;			
RL DR Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.			
DR DR EMBL; Z70718; CAA94674.1; -.			
DR DR EMBL; Z68222; CAA92502.1; -.			
DR DR EMBL; Z70284; CAA94278.1; -.			
DR DR EMBL; Z70284; CAA94283.1; -.			
DR DR HSSP; P27439; 3MSP.			
DR DR InterPro; IPR000535; MSP_domain.			
DR DR Pfam; PF00635; MSP_domain; 1.			
SQ SQ SEQUENCE 127 AA; 14236 MW; E5B96631BEBF1419 CRC64;			
RT "Genome sequence of the nematode <i>C.elegans</i> : A platform for investigating biology.";			
RT Science 282:2012-2018(1998).			
RL EMBL; Z81122; CAB03362.1; -.			
RN HSSP; P27439; 3MSP.			
DR InterPro; IPR000535; MSP_domain.			
DR Pfam; PF00635; MSP_domain; 1.			
DR DR Sequence 127 AA; 14233 MW; 69F69622069E2147 CRC64;			
RA none;			
RT "Genome sequence of the nematode <i>C.elegans</i> : A platform for investigating biology.";			
RT Science 282:2012-2018(1998).			
RL EMBL; Z81122; CAB03362.1; -.			
RN HSSP; P27439; 3MSP.			
DR InterPro; IPR000535; MSP_domain.			
DR Pfam; PF00635; MSP_domain; 1.			
DR DR Sequence 127 AA; 14233 MW; 69F69622069E2147 CRC64;			
Query Match 71.4%; Score 90; DB 5; Length 127;			
Best Local Similarity 100.0%; Pred. No. 4.4e-89;			
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
[1] Q27280 PRELIMINARY; PRT; 127 AA.			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Created)			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
ID Q27280; 01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
ID Q27280; NCBITAXID=6239; HYPOTHETICAL 14.2 KDA PROTEIN.			
RESULT 6			
ID Q27280 PRELIMINARY; PRT; 127 AA.			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Created)			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
ID Q27280; NCBITAXID=6239; HYPOTHETICAL 14.2 KDA PROTEIN.			
Query Match 71.4%; Score 90; DB 5; Length 127;			
Best Local Similarity 100.0%; Pred. No. 4.4e-89;			
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
[1] Q27280 PRELIMINARY; PRT; 127 AA.			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Created)			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
ID Q27280; 01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
ID Q27280; NCBITAXID=6239; HYPOTHETICAL 14.2 KDA PROTEIN.			
RESULT 6			
ID Q27280 PRELIMINARY; PRT; 127 AA.			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Created)			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
ID Q27280; NCBITAXID=6239; HYPOTHETICAL 14.2 KDA PROTEIN.			
Query Match 71.4%; Score 90; DB 5; Length 127;			
Best Local Similarity 100.0%; Pred. No. 4.4e-89;			
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
[1] Q27280 PRELIMINARY; PRT; 127 AA.			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Created)			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
ID Q27280; 01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
ID Q27280; NCBITAXID=6239; HYPOTHETICAL 14.2 KDA PROTEIN.			
RESULT 6			
ID Q27280 PRELIMINARY; PRT; 127 AA.			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Created)			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
ID Q27280; NCBITAXID=6239; HYPOTHETICAL 14.2 KDA PROTEIN.			
Query Match 71.4%; Score 90; DB 5; Length 127;			
Best Local Similarity 100.0%; Pred. No. 4.4e-89;			
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
[1] Q27280 PRELIMINARY; PRT; 127 AA.			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Created)			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
ID Q27280; 01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
ID Q27280; NCBITAXID=6239; HYPOTHETICAL 14.2 KDA PROTEIN.			
RESULT 6			
ID Q27280 PRELIMINARY; PRT; 127 AA.			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Created)			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
ID Q27280; NCBITAXID=6239; HYPOTHETICAL 14.2 KDA PROTEIN.			
Query Match 71.4%; Score 90; DB 5; Length 127;			
Best Local Similarity 100.0%; Pred. No. 4.4e-89;			
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
[1] Q27280 PRELIMINARY; PRT; 127 AA.			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Created)			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
ID Q27280; 01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
ID Q27280; NCBITAXID=6239; HYPOTHETICAL 14.2 KDA PROTEIN.			
RESULT 6			
ID Q27280 PRELIMINARY; PRT; 127 AA.			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Created)			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
ID Q27280; NCBITAXID=6239; HYPOTHETICAL 14.2 KDA PROTEIN.			
Query Match 71.4%; Score 90; DB 5; Length 127;			
Best Local Similarity 100.0%; Pred. No. 4.4e-89;			
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
[1] Q27280 PRELIMINARY; PRT; 127 AA.			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Created)			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
ID Q27280; 01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
ID Q27280; NCBITAXID=6239; HYPOTHETICAL 14.2 KDA PROTEIN.			
RESULT 6			
ID Q27280 PRELIMINARY; PRT; 127 AA.			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Created)			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
ID Q27280; NCBITAXID=6239; HYPOTHETICAL 14.2 KDA PROTEIN.			
Query Match 71.4%; Score 90; DB 5; Length 127;			
Best Local Similarity 100.0%; Pred. No. 4.4e-89;			
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
[1] Q27280 PRELIMINARY; PRT; 127 AA.			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Created)			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
ID Q27280; 01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
ID Q27280; NCBITAXID=6239; HYPOTHETICAL 14.2 KDA PROTEIN.			
RESULT 6			
ID Q27280 PRELIMINARY; PRT; 127 AA.			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Created)			
ID Q27280; 01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
ID Q27280; NCBITAXID=6239; HYPOTHETICAL 14.2 KDA PROTEIN.			
Query Match 71.4%; Score 90; DB 5; Length 127;			
Best Local Similarity 100.0%; Pred. No. 4.4e-89;			
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
[1] Q2			

RT	"The sequence of <i>C. elegans</i> cosmid C09B9.";
RL	Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN	[3]
RA	None;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RC	Wohldmann P.;
RA	"The sequence of <i>C. elegans</i> cosmid K05F1.";
RL	Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
RN	[3]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RC	Waterston R.;
RA	"Direct Submission.";
RL	Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; U29377; AAA68714.1; -.
DR	HSSP; P27439; 3MSP.
DR	InterPro; IPR00535; MSP_domain.
DR	Pfam; PF00635; MSP_domain; 1.
KW	Hypothetical Protein.
SQ	SEQUENCE 127 AA; 14228 MW; 7D5D628EF5511AA7 CRC64;
Query Match	56.38; Score 71; DB 5; Length 127;
Best Local Similarity	100.0%; Pred. No. 1.4e-68;
Matches	71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	31 IKVINSARRIGYGIKTTNMKRLGVDPGVLDPKEAVLLAVSCDAFAFGQEDTNNDRIT 90
Db	32 IKVINSARRIGYGIKTTNMKRLGVDPGVLDPKEAVLLAVSCDAFAFGQEDTNNDRIT 91
QY	91 VEWTNTPDGAA 101
Db	92 VEWTNTPDGAA 102
RESULT	8
ID	Q17856
ID	Q17856; PRELIMINARY; PRT; 127 AA.
AC	Q17856;
DT	01-NOV-1996 (TREMBLrel. 01, Created)
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE	C09B9.6 PROTEIN (HYPOTHETICAL PROTEIN R13H9.2).
GN	C09B9.6 OR R13H9.2.
OS	Caenorhabditis elegans.
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC	Rhabditidae; Pelerinae; Caenorhabditis.
NCBI_TaxID	6239;
[1]	SEQUENCE FROM N.A.
RP	STRAIN=BRISTOL N2;
RX	MEDLINE=94150718; PubMed=7906398;
RA	Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Bonfield J., Burton J., Connell M., Copsey T., Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten J., Laird N., Latreille P., Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R., Smalldon N., Smith A., Sonnhammer E., Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RA	"2.2 Mb of contiguous nucleotide sequence from chromosome III of <i>C. elegans</i> ." Nature 368:32-38(1994).
RT	SEQUENCE FROM N.A.
RC	STRAIN=BRISTOL N2;
RA	Bentley D.;
RT	"The sequence of <i>C. elegans</i> cosmid C34F11." Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
RL	[2]
RN	SEQUENCE FROM N.A.
RC	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[3]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	Bradshaw H.;
RT	"The sequence of <i>C. elegans</i> cosmid C34F11." Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
RL	[3]
RN	SEQUENCE FROM N.A.
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	Bradshaw H.;
RT	"The sequence of <i>C. elegans</i> cosmid C34F11." Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
RL	[3]
RN	SEQUENCE FROM N.A.
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
SEQUENCE FROM N.A.	
RP	STRAIN=BRISTOL N2;
RA	"Genome sequence of the nematode <i>C. elegans</i> : a platform for investigating biology. The <i>C. elegans</i> Sequencing Consortium.";
RL	Science 282:2012-2018(1998).

RA MSP.
 RT "The sequence of *C. elegans* cosmid ZK546.";
 RL Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; U29380; AAA68736.1; -.
 DR HSSP; P27439; 3MSP.
 DR InterPro; IPR00535; MSP_domain.
 DR Pfam; PF00635; MSP_domain; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 133 AA; 14899 MW; B283C4CDFDF9B76A CRC64;
 Query Match 40.5%; score 51; DB 5; Length 133;
 Best Local Similarity 100.0%; Pred. No. 5.6e-47;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 51 KRLGVDPGPCGVLDPEAVSCDAFAFGQEDTNNDRTITVEWTNTPDGAA 101
 DB 58 KRLGVDPGPCGVLDPEAVSCDAFAFGQEDTNNDRTITVEWTNTPDGAA 108
 RESULT 13
 Q9GNW0 PRELIMINARY; PRT; 107 AA.
 AC Q9GNW0;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MAJOR SPERM PROTEIN (FRAGMENT).
 GN MSP.
 OS Onchocerca volvulus.
 OC Eukaryota; Metazoa; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Onchocerca.
 OC NCBI_TAXID=6282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OVNODBRA2C3 ;
 RX MEDLINE=21321180; PubMed=11428336;
 RA Morales Hojas R.; Post R.J.;
 RT "Regional genetic variation in the major sperm protein genes of
 Onchocerca volvulus and *Mansonella ozzardi* (Nematoda: Filarioidea).";
 RL Int. J. Parasitol. 30:1459-1465(2000).
 DR EMBL; AJ404206; CAC20740.1; -.
 DR HSSP; P27439; 1MSP.
 DR InterPro; IPR00535; MSP_domain.
 DR Pfam; PF00635; MSP_domain; 1.
 FT NON_TER 1 107
 FT NON_TER 107 107
 SQ SEQUENCE 107 AA; 12018 MW; 0C5E932130E5DF07 CRC64;
 Query Match 17.5%; score 22; DB 5; Length 107;
 Best Local Similarity 100.0%; Pred. No. 9.2e-16;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 45 IKTTNMKRLGVDPGPCGVLDPE 66
 DB 39 IKTTNMKRLGVDPGPCGVLDPE 60
 RESULT 14
 Q9GNV9 PRELIMINARY; PRT; 107 AA.
 AC Q9GNV9;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MAJOR SPERM PROTEIN (FRAGMENT).
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OVNODBRA2C5 ;
 RX MEDLINE=21321180; PubMed=11428336;
 RA Morales Hojas R.; Post R.J.;
 RT "Regional genetic variation in the major sperm protein genes of
 Onchocerca volvulus and *Mansonella ozzardi* (Nematoda: Filarioidea).";
 RL Int. J. Parasitol. 30:1459-1465(2000).
 DR EMBL; AJ404208; CAC20742.1; -.
 DR HSSP; P27439; 1MSP.
 DR InterPro; IPR00535; MSP_domain.
 DR Pfam; PF00635; MSP_domain; 1.
 FT NON_TER 1 107
 FT NON_TER 107 107
 SQ SEQUENCE 107 AA; 11987 MW; 0C4325F7F6E5DF07 CRC64;
 Query Match 17.5%; score 22; DB 5; Length 107;
 Best Local Similarity 100.0%; Pred. No. 9.2e-16;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: September 24, 2002, 11:04:07
 Job time: 156 sec

